

Keeping faith in gene manipulation

There are good reasons to revise the way that gene therapy experiments receive public scrutiny. But altering the role of a body with a reputation for effective oversight must be done sensitively if it is not to undermine efforts to maintain public acceptance of the field.

VIRTUALLY from the day it was created in the mid-1970s, the Recombinant DNA Advisory Committee (RAC) of the US National Institutes of Health (NIH) has remained highly controversial within the scientific community. Initially, it was those who claimed that the potential health and environmental dangers of gene-splicing had been vastly overestimated by an imaginative public that led to demands for its dismantling, on the grounds that its activities imposed an unacceptable bureaucratic burden on researchers. More recently, the committee has turned its attention to the ethical dimensions of gene therapy. As it has done so, others have championed its abolition by arguing that the research raises few issues that cannot be adequately dealt with by more conventional regulatory bodies — and indeed that singling out gene therapy for special attention itself raises an unwarranted danger signal in the eyes of the public.

In both cases, critics of the RAC have been answered at two levels. One is that the promises — and threats — of genetic engineering are, thanks to the ability of recombinant DNA techniques to alter the essence of living organisms, qualitatively different from previous laboratory practices and clinical procedures. The second is that the RAC, while acting as a focal point of public concerns, has also helped to contain these concerns by providing a structured forum in which they can be safely debated. Indeed, it could be argued that the RAC's high-profile existence has helped to stave off demands for restrictive legislation whose existence has, in countries such as Germany, helped to stifle the birth of a nascent biotechnology industry.

Many may be tempted to dismiss the efforts of Harold Varmus, the director of NIH, to eliminate the RAC as a capitulation to the demands of abolitionists. It would be wrong to do so. Varmus's arguments, set out in a formal proposal published last week in the *Federal Register* (see page 195), are based on a careful analysis of both the strengths and weaknesses of the committee's activities. In essence, his case is that the RAC spends too much time studying the details of relatively uncontroversial experiments, and could focus its efforts more effectively on the broader health and ethical issues raised by novel applications of gene therapy, such as its use on unborn children, which may amount to inadvertent germ-line therapy.

To achieve this, Varmus is proposing a clear separation between regulatory and 'monitoring' responsibilities. The first would be transferred entirely to the Food and Drug Administration (FDA) (where much of it already resides). The second would remain at NIH, but would be passed to a new, more streamlined body, with the ungainly title of the Office of Recombinant DNA Activities Advisory Committee (OAC). In addition to regular meetings of its own, the latter would be responsible for identifying topics and participants for regular public conferences, each dedicated to a single topic related to the science and ethics of novel techniques and their clinical applications.

The attractions of this restructuring are obvious. For gene-therapy researchers, it will provide a less burdensome regulatory framework. For the NIH, it will permit public discussion of novel research without, in its own words, having to give "a stamp of approval on the basis of a limited threat to human health or

safety". The difficulty arises over the question of public scrutiny. One of the merits of the RAC has in the past been the way in which open discussion of individual experiments has given them a particularly robust form of public legitimation. Conducting the same debates within the privacy of FDA committees will not be the same.

The key issue at stake with gene therapy, over and above the relative merits of individual experiments or techniques, is the maintenance of public trust. There was a time when this could be sustained simply by following the advice of panels of scientific and ethical experts. But the world has become more sophisticated. The certainties from which expert advice is offered are no longer taken for granted; regulatory justice must not only be done — it must be seen to be done.

Ways of achieving this in practice inevitably vary from one country to another, depending on indigenous political and cultural traditions. France, for example, has tended to take a legislative approach, enshrined in the decision to give cell and gene therapies the status of pharmaceutical drugs under the *Agence de médicament* — the French equivalent of the FDA — while requiring their use to comply with laws on the release of genetically modified organisms; excessive rigidity is said to be avoided by the informal flexibility with which laws are traditionally interpreted on the ground.

In Germany, the situation is similar, except that the regulations have in the past been more demanding — reflecting sensitivities heightened by the genetic excesses of the Nazi era — while their flexible interpretation is less widely accepted. Japan, also carrying a historical legacy, has felt the need to make significant efforts to secure public acceptability, with academic researchers facing two parallel sets of regulatory hurdles, one on behalf of the Ministry of Health and Welfare, the other of the Ministry of Education, Science and Culture. Britain has taken a less burdensome approach. But again the expected level of vigilance has been high, first through the Genetic Manipulation Advisory Group, most recently with the government's agreement, under pressure from parliament, to set up a broad-ranging Human Genetics Advisory Commission.

In the United States, the concern that appears to unite the critics of moves to eliminate the RAC is that removing decisions for novel gene-therapy protocols from direct public scrutiny threatens to undermine the very process by which public legitimacy is traditionally created.

Varmus's proposals indicate that he is far from deaf to these concerns. But it remains important that they are respected in whatever machinery evolves after the present consultation period. One requirement is that links between the FDA and the NIH's remaining oversight machinery should remain strong and transparent, ensuring that all initiatives requiring public scrutiny continue to receive it. A related suggestion is that the new body, rather than being represented as a clear break with the past, should be re-cast as the product of an evolutionary process. Perhaps keeping the name of the RAC, and, more importantly, providing some guarantees that its opinions will continue to wield influence, would reassure a sceptical public that the changes are for the best. □



Satellite missions face shut-down as NASA tackles cash squeeze

Washington. A panel of space scientists is to meet later this month to help the US National Aeronautics and Space Administration (NASA) to decide which astrophysics satellites should have their lifetimes extended and which may have to be shut down for lack of money.

Although NASA managers conduct a similar exercise every two years, this year's 'senior review' by scientists from outside the agency promises to be the toughest yet. The agency's budget for mission operations continues to decline, while the number of spacecraft returning data beyond their planned lifetimes is growing.

The panel will judge the scientific merit of extending three US missions scheduled to end in 1997 and 1998: the Extreme Ultraviolet Explorer (EUVE), the Compton Gamma Ray Observatory (GRO) and the Rossi X-Ray Timing Explorer (XTE).

Also to be decided is continued US participation in several international missions: Japan's ASCA X-ray satellite, the German-led ROSAT X-ray mission, Italy's Beppo-SAX, Japan's VLBI Space Observatory Programme (VSOP), and grant support for US users of Europe's Infrared Space Observatory (ISO).

These missions will cost NASA a total of \$38 million next year. This should present no difficulty, says Guenter Riegler of NASA's science office. The problem comes in 2000, when the budget for these projects will go down to \$6 million. "Clearly something has to give," says Riegler, adding that the agency may need to terminate some missions earlier than others to save money for later in the decade.

After hearing from the review panel, NASA will decide next month which spacecraft, if any, should be shut down, and when. The agency plans to conduct similar exercises for its space physics and planetary programmes within the next two years.

No one inside or outside NASA relishes throwing away a working satellite. But the agency did just that by deciding, on the advice of previous review panels, to abandon the ailing but still productive International Ultraviolet Explorer (IUE), which will shut down in September after a brief period of European ownership.

The IUE decision was "a big psychological barrier", says Riegler. He says that such painful decisions may become increasingly common as the space agency's operational budget dwindles.

While NASA may budget only for a two- or three-year mission, spacecraft routinely

last longer, returning useful data for five or more years. The problem is exacerbated by NASA's current philosophy of launching more small satellites, more often.

"The thinking now seems to be 'better, cheaper, faster, then throw it away'," says one scientist, who argues that investigators often need a year or two to work with new spacecraft instruments before they get the most out of them.

Many scientists are also concerned that the squeeze on budgets means that more money is being spent on keeping hardware

IMAGE
UNIMAGEABLE
FOR A COYALBET
FOR REASON
REASONS

NASA

On the bright side: scientific merit is likely to help the Gamma Ray Observatory survive.

flying than on supporting science. Project scientists already report that funding for data analysis has shrunk disproportionately, with fewer researchers getting ever smaller grants to analyse spacecraft results.

If NASA wants to continue launching missions with new sensors, it cannot afford to keep older spacecraft flying indefinitely. Most space scientists acknowledge that this is a genuine dilemma in an era of declining budgets. "NASA really does want to do the right thing," says Fred Seward, a member of the ROSAT team at the Harvard-Smithsonian Astrophysics Observatory. He laments: "I feel like we're at war, and we're trying to defend ten cities, and we're told we can only defend five — which ones do we want?"

Continued US participation in ROSAT and ASCA may be hard to justify this year, particularly with a small armada of X-ray satellites already flying, and the US Advanced X-Ray Astrophysics Facility and European XMM missions scheduled for launch in 1998 and 1999 respectively.

Extending the EUVE seems to be virtually assured, however, as NASA plans to use it as an experiment in reducing operational costs by turning the satellite over to a university to run. The University of California at Berkeley, which already operates EUVE's science instruments from its Center for Extreme Ultraviolet Astrophysics (CEA),

would take over full responsibility for the satellite from NASA's Goddard Space Flight Center in Greenbelt, Maryland, perhaps as early as next winter.

Roger Malina, the director of CEA, says that, by relying heavily on artificial intelligence and other automated techniques, and by using students rather than government workers and contractors, Berkeley should be able to run the satellite "a lot cheaper" than NASA — as little as \$3.5 million a year.

Riegler is more cautious, saying that the plan to hand over EUVE is an "experiment" that will initially cost his office more money, at least in the transition phase when Goddard and Berkeley will share control of the spacecraft. But NASA's hope — and its policy with future scientific spacecraft — is to save money by "outsourcing" and automating mission operations as much as possible.

In the case of the larger and more complex GRO, outsourcing mission operations to a single university is unlikely. So the project has tried to save costs in other ways — for example, by switching from a multi-mission control room at Goddard to control through distributed workstations.

Such streamlining may eventually be able to cut operations costs in half, down to an estimated \$8 million a year by the end of the decade. But that alone would break NASA's allotted budget for fiscal year 2000 for operating the missions under consideration at this month's review.

And GRO, too, is likely to survive on scientific merit. It covers a wavelength range that no other satellite does, it addresses a high-priority scientific problem — the nature of gamma-ray bursts — and no other spacecraft is being built to replace it. "We don't have anyone nipping at our heels," says Gerald Fishman of NASA's Marshall Space Flight Center in Huntsville, Alabama, principal investigator for one of GRO's instruments.

Tony Reichhardt

Internet data warning

Paris. French research organizations and universities that include details of researchers in directories published on the Internet should first ask individual researchers for their formal authorization, according to CNIL, the body that supervises the enforcement of the French data-protection act. In its annual report, released last week, CNIL also warned that computer systems should be designed to prevent confidential medical data on individuals from falling into the wrong hands. (See also page 198.) □

Russian parliament's funding plans run into budget realities

Moscow. The State Duma, the lower chamber of the Russian parliament, has voted almost unanimously to increase the national budget for science and education next year by 15,250 billion roubles (US\$3.05 billion) — a rise of 60 per cent. But the rise is certain to be rejected by the parliament's upper chamber and President Boris Yeltsin.

The rise was put forward in an amendment to the budget bill by the committee on science and education, headed by Ivan Melnikov of the Communist Party. Oleg Smolin, deputy chairman of the committee, suggested that the extra money could be taken from companies "that enjoy privileges and favourable conditions".

But the Duma's budget committee objected to the amendment because only 70 per cent of the income for the overall federal budget could be met. Oksana Dmitrieva, chairwoman of the budget committee, called the amendment "the budget Bacchanalia". She said it was a populist gesture, like many made during the recent presidential election campaign, and that it was illegal to allocate increased financing without saying where the funds would come from.

Gennadi Kulik (Agrarian Party) and Yuri Voronin (Communist), both deputy chairmen of the budget committee, reminded the deputies that the law concerned had never been violated, and that it allowed the government to limit inflation. The Duma should either alter the law or follow it, they said. If no additional money could be found, the amendment would increase the budget deficit.

But the Speaker of the Duma, Gennadi Seleznev (Communist), dismissed these warnings as "formalities", while one Communist deputy said that, if it had not been for the expenditure on presidential elections, the money for science and education could easily have been found.

The budget committee suggested ways to provide substantial support for science and education without violating the budget. For example, it proposed that funds for science and education should be included in the 'protected' areas of the budget, to be met under any circumstances; that the government should compensate scientists and educationalists for losses suffered since it failed to increase their salaries as promised by its decree at the end of last year; and that expenditure should be indexed each quarter of the year if inflation proved higher than anticipated.

But the deputies rejected all but the last of these measures. They approved the amendment suggested by the committee on science and education by 235 to 2, with 17 abstentions.

Carl Levitin

Social needs proposed as priority for EU research

Munich. The European Union (EU)'s Fifth Framework Programme for research (FP5), to be launched in 1999, should focus more on social needs than the current programme (FP4), according to a discussion paper prepared under the auspices of Edith Cresson, the research commissioner.

The paper, which was approved last week by the full commission and presented to the Council of Ministers, says that FP5 should be organized around a limited set of objectives, ensuring that these meet the social and economic needs of European citizens.

This indicates a shift in the commission's approach to its research programmes, which currently cost an average of ECU2.6 billion (US\$3.2 billion) a year, and which have until recently been dominated by industrial needs. Although these needs remain fundamental, the commission is now also stressing that research should have a direct impact on people's lives.

The discussion document is being widely distributed to member states and basic and industrial research organizations. According to one of its co-authors, Jack Matthey, a policy expert in Cresson's office, it will "be used as a basis for wide reflection" before the commission makes its final proposal on the content and operation of FP5 to the Council of Ministers and the European Parliament next spring.

The three broad research areas that the commission proposes throw up few surprises. The first, under the heading "Unlocking the resources of the living world and the ecosystem", would include "acquisition and utilization of knowledge about fundamental mechanisms affecting human life, especially in the fields of health and food".

The second, "Creating a user-friendly information system", covers aspects of advanced communication technologies that meet the needs of industry and society. Priorities would include mechanisms for facilitating access to information and helping to preserve linguistic diversity.

The third area is "Promoting competitive and sustainable growth". This would cover industrial production methods that take into account the full life-cycle of a product — including waste management — as well as the management of energy, transport, agriculture and forestry.

The paper also proposes three additional activities to accompany the main research

programmes. "Improving human potential" would cover training and mobility, the promotion of a European science prize, and a socio-economic programme to analyse the full social impact of research activities. The second activity would be designed to increase the participation of small businesses in FP5, for example by tailoring procedures more closely to their needs and by helping to establish links with venture capital companies.

The third activity would be to foster international links with European research, not only with non-EU industrial countries and the emerging economies of Asia, but also with central and eastern European countries that are expected to join the union in the near future.

According to Matthey, the paper is intended primarily to indicate the commission's general approach to FP5. Its eventual proposals on the content of the research will be based on discussions with interested parties over the next few months, he says. "We will continue to consider not only the industrial point of view, but also the environmental and social points of view."

Under the rules of the Maastricht Treaty, the Framework programme requires the unanimous vote of the EU's 15 member states and majority approval by the European Parliament. Commission officials argue that early consensus is the only pragmatic way forward. They also acknowledge that the commission's perceived lack of transparency and the inflexibility of its programmes are widely criticized; the document makes improved user-friendliness a priority.

The commission has in fact already made significant moves in this direction, for example by publishing details of its research programmes more widely, and simplifying application procedures. Ironically, this is held partly responsible for the heavy over-subscription of Framework programmes.

Further proposals aimed at improving the management of research programmes include the development of more immediate monitoring procedures, speeding up the selection of projects and simplifying the procedures involved in implementing programmes — by reducing the number of committees involved, for example.

The commission wants to encourage greater cooperation between member states by using more of the funding mechanisms available under the Maastricht Treaty. It would, for example, like to be able to supplement some activities of national research organizations, and to restrict some FP5 programmes to only a few member states, an approach known as 'variable geometry'.

Alison Abbott



Cresson: FP5 should improve people's lives.

Budget cuts threaten Israeli science jobs

Jerusalem. Government-funded research jobs for immigrant scientists from the former Soviet Union may fall victim to the new Israeli government's determination to slash its budget, according to scientists lobbying for subsidies to fund such posts.

Last week, the cabinet rejected a proposal from the Finance Ministry to eliminate funding for programmes that support immigrants working on university and private-sector research teams. But it stipulated that any expansion of these programmes would have to come out of the existing budget of the Ministry of Absorption, which is responsible for accommodating immigrants from the Soviet Union.

The cabinet also decided to set up a committee to decide whether current and proposed science job-creation programmes are "essential" and "worthwhile". A spokesman for Yuli Edelstein, the Minister of Absorption, called the cabinet decision a "victory", as it preserved US\$1 million in job subsidies approved before the election by the previous Labor government. Both Edelstein and Natan Scharansky, the Minister of Industry

and Commerce, belong to Yisra'el ba-Aliya, a political party representing the interests of immigrants from the former Soviet Union.

But reactions from Israeli scientists who support the subsidies have ranged from puzzlement to exasperation. The coalition agreement signed by Yisra'el ba-Aliya with the Likud party of the prime minister, Benjamin Netanyahu, explicitly promises funding for existing programmes and to examine the possibility of expanding them.

Nathan Patlas, assistant professor of anatomy at the Hadassah Medical School of the Hebrew University of Jerusalem, and coordinator of immigrant scientists for Yisra'el ba-Aliya, says he has been told by Edelstein and Scharansky that the government's budget cuts meant that there would be no money for subsidies. "The immigrant scientists are very worried," he says.

Jobs for scientists became a political issue as a result of the mass immigration that occurred at the beginning of the decade, when more than half a million people arrived in Israel from the Soviet Union. Among them were 12,000 scientists — more than Israel itself had produced since independence in 1947.

The universities offered faculty positions to 180 of the top-flight newcomers. About 2,000 others were accepted, over the course of several years, joining research teams led by Israeli academics. The Ministry of Absorption provided minimal salaries for three years, hoping the scientists would then be able to find jobs in the private sector.

But not all university officials support the

subsidies. Amnon Pazy, professor of mathematics at the Hebrew University of Jerusalem, and chairman of the Planning and Grants Commission, which allocates funding to universities, fears that the costs will be passed on to the universities, and is concerned that the immigrants might end up as cheap research assistants.

Even with the help of these programmes, claims Yakov Faitelson, director-general of the Israel Association of Immigrant Scientists, 70 per cent of the immigrant researchers were left without work, or had to work outside their profession. Over the past year, as the three-year grants began to run out, the immigrants faced the loss of their jobs.

Dan Huppert, professor of physical chemistry, and three colleagues at Tel Aviv University, proposed to the government that it should subsidize lifetime positions for 400 of the most valued researchers. This idea was adopted last May by the government of the former prime minister, Shimon Peres.

Huppert's group also proposed programmes that would subsidize many more jobs in the academic and private sectors. In the wake of last May's elections, he and his colleagues had hoped that Yisra'el ba-Aliya would use its political clout in the new government to get these programmes funded.

But with Netanyahu determined to slash the national budget, it is not clear that the government will continue to fund even the existing programmes beyond this year.

Haim Watzman

Private universities in Japan win grant boost

Tokyo. Twenty-two of Japan's private universities are to receive generous "high-tech research centre" grants worth an average of about ¥1 billion (US\$9 million) under a new government scheme.

Apart from a handful of outstanding institutions, such as Waseda, Keio and Tokai universities, Japan's 410 private universities have not developed strong reputations for research, partly because they do not receive the same level of financial support for such work as the 98 national universities.

The new programme is intended to help redress the imbalance. The grant-winners were selected from 78 applications from 55 institutions. Half of the grants will go towards biological and medical research, and the others will support projects in information science, engineering and other sciences, with the emphasis on basic rather than applied research.

The Ministry of Education, Science and Culture has set aside ¥4.7 billion in the current fiscal year, which began on 1 April, for buildings and facilities for the programme. This sum will be supplemented by additional grants from the ministry and the Japan Society for Promotion of Science, which will require matching funds from the private universities' own resources.

Details of the research programmes have yet to be mapped out; a ministry committee selected the winning applications merely on the basis of a brief summary of proposed research.

Stephen Barker

Heat turned up on climate author

Washington. Benjamin Santer, the lead author of a controversial chapter of the Second Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), and a climate scientist at the Lawrence Livermore National Laboratory in California, has come under attack by a prominent US congressman for his role in modifying the chapter's contents.

Dana Rohrabacher, the chair of the energy and environment subcommittee of the House of Representatives Science Committee, has written to Hazel O'Leary, the energy secretary, demanding information about Department of Energy (DOE) funding for the work of Santer and all other IPCC authors. He also wants the DOE to regulate their work for IPCC.

The letter repeats allegations made by the Global Climate Coalition (GCC), an industry-funded pressure group, that Santer broke the IPCC's own rules by modifying the content of Chapter 8 of the Second Assessment Report after it

had been accepted by IPCC's science working group at Madrid last November (see *Nature* **381**, 546; 1996).

Rohrabacher is a strident critic of the theory that global climate change is caused by man, which he has called "liberal claptrap". His subcommittee does not have direct power over research budgets, but it can influence policy by scrutinizing programmes it dislikes. In this case, the letter will send DOE officials on a lengthy paper chase after evidence to back up the work being done at its laboratories to support the IPCC process.

Sir John Houghton, formerly chief executive of the UK Meteorological Office and currently chairman of the Royal Commission on Environmental Pollution, who is co-chair of the IPCC science working group, backs Santer against this latest attack. "The allegations made are completely without foundation," he says.

Colin Macilwain

German plans may increase competition ...

Munich. The German cabinet last week approved plans drawn up by the research ministry to improve the efficiency of federally funded research institutes. Jürgen Rüttgers, the research minister, said that the proposed changes — most of which have to be approved by the *Länder* (state) governments — would “allow us to fund the best [research] through competition”.

One change is that scientists in so-called ‘blue list’ institutes will compete with university scientists for part of the money that they at present receive automatically.

Scientists have welcomed the plan, but still express concern in private. Initially the research ministry had been thinking of transferring up to 20 per cent of the blue list’s institutional financing to the Deutsche Forschungsgemeinschaft (DFG), Germany’s basic research funding agency, to be redistributed as project money through open competition. But, aware that the plan needs to be introduced cautiously if it is to win public acceptance, this transfer will be limited to 5 per cent under the plan approved by the cabinet.

Scientists are now hoping that the ministry will also be prepared to compromise on its desire that all blue-list research institutes — including those in the east, which have hardly begun to find their feet — will be included in its plan.

Blue-list institutes are a mixed bag of research and service institutes jointly funded by federal and *Länder* governments. After reunification, many institutes of the East German Academy of Sciences that were judged by the German science council, the Wissenschaftsrat, to be of good quality, were placed in this category (see *Nature* 363, 482; 1993). This nearly doubled their total number, now standing at 83. But the increased funding involved attracted the attention of critics, who argued that such institutes had never been subject to systematic quality control (see *Nature* 377, 466; 1995).

According to Ingo Hertel, president of the WBL, the new organization that defends the general interests of the blue-list institutes, the proposal to put part of their budgets into a common pot of project money will increase the amount of quality

control and raise the level of research. But he says that its success will depend on how the idea is put into practice. It should not be imposed on all institutes, he says.

Christof Schneider, a spokesman for the DFG, agrees. Blue-list institutes may object if they fail to win back the research funds on which they used to rely, and see the money going to universities. Some of the new institutes in the east have not yet even appointed directors, he says, and are too young to have to face such a strain.

The DFG has other concerns about the plan. Despite an annual increase in funding of 5 per cent over the past few years, the agency is struggling with an increase in the number of research applications, mostly from universities, which has grown at the much faster rate of 10 per cent per year. The additional burden of handling even more applications from blue-list scientists could prove too great to maintain the efficiency that the research ministry praised in its paper.

Second, the DFG is concerned about what Schneider calls the “ethical question” of whether giving it responsibility for part of the blue-list budget could force the agency into an unwelcome political role that it has successfully avoided in the past.

Negotiations are taking place to resolve these issues. But it is already clear that the hopes of the research ministry that the new system will be in place by next January are unrealistic. Any changes to the budgetary system of blue-list institutes requires the formal approval of the *Länder* governments, and negotiations are likely to be difficult.

The research ministry’s strategy paper also suggests that parts of the budgets of Germany’s 16 national research centres, often criticized as inefficient (see *Nature* 372, 4; 1994), should eventually be handled by the DFG in the same way.

More immediately, it suggests that the newly created senate of the Helmholtz Society, which represents the centres nationally, should control part of the general institutional budgets, which cover running costs and salaries, directing the money to national research centre departments working in priority areas which it will be responsible for defining. The senate comprises experts from the political, scientific and industrial areas.

Joachim Treusch, who is director of both the Helmholtz Society and the KFA research centre in Jülich, is enthusiastic about this suggestion, which he believes will help the national research centres to contribute effectively to Germany’s strategic needs. But he does not support a future role for the DFG in handling project money. That would make the financing of the centres too complicated, he says, if they already have to apply for money from the senate and from industry.

Alison Abbott & Quirin Schlermeier

...and hammer geological research

Munich. German scientists, who until recently have escaped the significant budget restrictions experienced by their European neighbours, will soon be sharing the same fate. Jürgen Rüttgers, Germany’s research minister, announced last week that his ministry is not to be exempt from cuts averaging 2.5 per cent being applied to all government spending in the federal budget for 1997.

The latest cuts are intended to help Germany meet the financial criteria for entry into the European monetary union. When added to budget cuts announced earlier this year (see *Nature* 381, 9; 1996), the result will be a reduction of 4.5 per cent next year on the originally-approved 1996 budget for science and research.

Geological sciences will suffer most in financial terms, with a 40 per cent reduction in funding, although much of this will be the result of the completion of Germany’s large continental drilling programme in northern Bavaria, known as KTB (Kontinentaltieftiefbohrprogramm), at the end of this year.

Rolf Emmermann, the scientific director of the Geoforschungszentrum in

Potsdam, Germany’s largest geological sciences institute, says he hopes the research ministry will eventually agree to support a replacement large-scale Earth sciences programme. But he accepts that this is no longer guaranteed.

“We will have to explain our relevance to society,” says Emmermann. He predicts that there will be a move away from basic research to “some kinds of early-warning systems for earthquakes, floods, or volcanoes”. The whole of Earth sciences is entering a crisis, he says, as nearly all publicly funded research projects will run out by 1998.

Ocean and polar research will also to be hard hit by the latest round of cuts, each being cut by 11 per cent in project money, and a smaller amount in institutional support.

In contrast, there is good news both for the Max Planck Society and the Deutsche Forschungsgemeinschaft (DFG), which funds university research. Government promises to maintain the annual 5 per cent increase they have enjoyed for the past six years will be honoured in the 1997 budget. **Q. S. & A. A.**

Going down: the world’s largest on-shore drill rig will soon finish its work.

Xenotransplantation trials 'should proceed but need guidelines'

Washington. A panel set up by the US Institute of Medicine (IOM) has concluded that the potential benefits of xenotransplantation — the transplantation of organs, cells or tissues from animals to humans — outweigh its risks. It therefore recommends that scientifically justified clinical trials of such transplants should go ahead, provided federal guidelines and other safeguards are put in place.

The recommendation comes in a report released this week by the IOM's Committee on Xenograft Transplantation, established in 1994 in response to increasing ethical and public health questions about xenotransplantation.

"When the science base for specific types of xenotransplants is judged sufficient and the appropriate safeguards are in place, well-chosen human xenotransplantation trials using animal cells, tissues and organs would be justified and should proceed," the report says.

Experiments with xenotransplantation of organs have not been carried out in the United States since the early 1990s. But several proposals are now under review, including some involving tissue transplantation to patients with AIDS and Parkinson's disease.

A key recommendation of the committee is that all institutions and individuals carrying out xenotransplants should adhere to national standards. The Food and Drug Administration and the Centers for Disease Control and Prevention are developing such guidelines, which are expected to be released soon.

The committee rejected proposals for a national body, such as the Recombinant DNA Advisory Committee of the National Institutes of Health, to oversee experimenters and ensure that they comply with new federal standards. Instead, it said, "appropriate" supervision could be provided by increasing the resources of local Institutional Review Boards and animal care committees, and requiring them to adhere to the guidelines in evaluating specific protocols.

But it does recommend that the Department of Health and Human Services should establish an advisory committee to "coordinate but not to regulate" research, policy and surveillance issues.

The committee also points out that "ample evidence" for the transmission of infectious agents from animals to humans indicates a danger "unequivocally greater than zero" that xenotransplantation could introduce new and deadly viruses into humans. **M. W.**

US plan to scrap DNA panel splits genetic researchers

Washington. US gene-therapy researchers and others appear to be split over a proposal to eliminate the Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH), set up in the mid-1970s in response to public fears of genetic engineering.

In a notice published in the *Federal Register* last week, the NIH describes as "timely and appropriate" a plan made public in May by Harold Varmus, the director of NIH, to disband the 25-member panel of scientists, ethicists and others, now responsible for approving novel gene-therapy protocols, and to transfer this task to the Food and Drug Administration (FDA).

But the proposal, on which the public has until 7 August to comment, has been criticized by politicians and scientists on the grounds that it will remove the possibility of public scrutiny of applications for experiments. (The FDA conducts reviews and approval meetings privately.)

In arguing for the RAC's dissolution, NIH cites the lack of novel protocols received under a new system of 'consolidated' review, initiated last year, under which unoriginal protocols, which make up 90 per cent of those received, are sent immediately to the FDA (see *Nature* 379, 668; 1996 & 375, 713; 1995).

NIH also argues that public airing of issues would not be lost. "It is not the RAC per se that is critical for public accountability, but the system by which NIH continues to provide public discussion," the *Federal Register* notice states.

Under the new plan, the NIH would retain a smaller committee of six to ten members, to be known as the Office of Recombinant DNA Activities Advisory Committee (OAC). This would meet 'regularly', and would organize public conferences on new and controversial gene-therapy topics up to four times a year, but would not review or approve specific protocols.

Many scientists are delighted at Varmus's bid to eliminate what they see as an unwarranted bureaucratic burden. Paul Berg of Stanford University, one of the scientists whose concerns in the early 1970s led to the setting up of the RAC, says: "The RAC has outlived its usefulness, and I don't see that it ever had any special competence for reviewing proposals for clinical trials." Although the proposed OAC would foster "very public

discussions", it "would not be sitting in the approval pathway for proposals".

Ronald Crystal, a gene therapist at the New York Hospital-Cornell Medical Center and the first gene therapist to use adenovirus for *in vivo* gene therapy, says that as a clinical investigator he found RAC approval "comforting" in earlier days. But, he adds, "I don't think it's necessary any more. The newness of gene therapy is no longer there."

Others are concerned about the proposal. "The proposed quasi-RAC is just a whitewash, just window dressing," says W. French Anderson, a molecular geneticist at the University of Southern California, who pioneered *ex vivo* human gene therapy in 1990 by treating a girl lacking adenosine deaminase. Unlike the RAC, the new committee, he says, will not be able to "critically analyze protocols as they come in".

Anderson complains that the proposed conferences on novel topics, which would cover policy aspects of topics such as *in utero* gene therapy, would not allow the public scrutiny of specific proposals that is assured by the current system.

His concern is shared by Malcolm Brenner, director of the cell and gene-therapy programme at St Jude's Children's Research Hospital in Memphis, Tennessee, and one of the earliest applicants to the RAC. Varmus's proposed conferences are "not going to get down to the nitty-gritty" of specific protocols, says Brenner.

He maintains that, whereas the RAC "compelled" researchers to discuss their plans publicly, "the FDA certainly won't do that". But Philip Noguchi, director of the FDA's cellular and gene therapies division, insists that "public oversight is still going to be there".

Another critic is Sheldon Krinsky, a former member of RAC and a science policy analyst at Tufts University in Medford, Massachusetts. Krinsky claims that the spectre of genetic manipulation for cosmetic ends opens an ethical Pandora's box. "We may be short-cutting the type of public dialogue that would be necessary for sound public policy," he says.

There is also disquiet in Congress, where some politicians have noted with alarm the proposal to disband the RAC. The subcommittee of the Senate appropriations committee that controls the NIH's \$12-billion or so budget last month asked Varmus for written justification of the proposal.

But it is not clear that Congress as a whole will rally to the RAC's defence. Senator Edward Kennedy (Democrat, Massachusetts), for example, often outspoken on such issues, has so far declined to take a stand.

Meredith Wadman



Varmus: plans smaller advisory committee.

Faraday attracts more interest in plans for UK science centres

London. Britain's Labour Party has promised to create a number of 'Faraday Centres' — intended to link the academic and industrial worlds, and named, ironically, after Mrs Margaret Thatcher's favourite scientist, Michael Faraday — if it wins the general election due within the next year.

The party's goal, according to Adam Ingram, the 'shadow' minister for science and technology, is to create "stable institutional structures to facilitate bringing together the various strands of our science base".

But despite a pledge in its manifesto, published two weeks ago, to recognize the importance of science to the economy, the Labour Party has held back from any commitment to increase funding for science — as it has for virtually all other areas of public spending.

The idea of creating Faraday Centres was first proposed by the Conservatives just before the last general election in 1992, based on a suggestion by a working party set up by Prince Charles that Britain should develop institutions similar in function to Germany's Fraunhofer Institutes.

Soon after, the Labour Party came up with its rival project for so-called 'Newton Centres'. But after the Conservatives won the election, the Faraday concept was dropped by the then president of the Board of Trade, Michael Heseltine.

In recent months, the Conservative government has again been promising to set up a small number of centres, aimed at providing focal points for interaction between the private sector and the publicly-funded research community.

Labour now seems keen on a similar idea. It has dropped its earlier insistence that these be new, free-standing institutions, and accepted that they should be based on existing research centres. Ingram says that Labour is keen to use public sector research establishments as a nucleus.

He also says that he has found a high level of enthusiasm in private industry, spurred in particular by the government's Technology Foresight exercise. "There seems to be plenty of interest from the private sector in fields such as aeronautics, space and marine research, materials, the chemical industry and pharmaceutical research," he says.

David Dickson



Faraday: name used by political rivals.

Network crime prompts UK moves on access to data

London. The British government is planning to give itself powers to access encrypted information stored on — or passing through — computer networks, according to proposals set out last month by the Department of Trade and Industry (DTI).

The move follows concern among Britain's intelligence agencies over a perceived increased use by criminals of sophisticated encryption software to conceal information used in unlawful activities such as computer pornography and fraud.

It also coincides with a renewed attempt by the United States government to introduce the so-called 'key escrow' system, after earlier moves in this direction had been dropped. A cabinet-level committee is to propose new encryption recommendations to President Bill Clinton by early September.

But the UK proposals are being viewed with concern by professional bodies such as the British Medical Association (BMA), which highlights the potential loss of privacy for lawful users of encrypted public networks, such as the National Health Service.

If the proposals become law, any individual or organization intending to send or store encrypted information on a network linked to an external telephone line will be required to provide copies of the decryption passwords — known as keys — to licensed independent organizations, known as 'trusted third parties'.

Government departments and intelligence agencies will — on production of a court order to the appropriate third party — be allowed access to these keys if they need to monitor encrypted information. This system is referred to as 'key escrow'.

The scheme is in line with identical proposals put forward by the European Commission (EC) last August (see *Nature* 377, 275; 1995). Parts of it are also similar to key escrow proposals put forward in the United States by the Clinton administration in 1993, but strongly opposed by civil liberties organizations and business groups.

Although civil liberties groups in the United Kingdom are similarly opposed to the British proposals, the government has been careful not to irritate the business community. The DTI scheme, for example, will not permit access to encrypted information on private-company mailing lists or in closed computer networks such as those between banks and their customers. "It is not the intention of the government to stifle competitiveness," says a DTI spokesman.

But not all potential critics have been satisfied. The BMA is deeply opposed to the idea of government access to encrypted data. The proposals have fuelled a dispute between the BMA and the Department of

Health about the security of a planned computer network for the health service which will allow networked storage and retrieval of patient records.

Criticism has also focused on suggestions that the General Medical Council (GMC), the doctors' regulatory body, could act as the third party to release decryption keys.

Ross Anderson, a lecturer at the University of Cambridge Computer Laboratory and an adviser to the BMA on information security, says the proposal to ask the GMC to hold decryption keys would transform it from a body charged with upholding professional standards into one responsible for disclosing personal health information to the security services. "I cannot imagine a proposal more likely to destroy the high level of trust that the [council] currently enjoys from both professionals and patients," he says.

Gunnar Klein, who chairs an EC working party drawing up a common standard for European health-care information networks, says he does not see how government access to confidential patient records will help to prevent the passage of information used in illegal activities.

Klein finds the UK government's regulatory approach inconsistent. The proposals, he says, do not give the government the right to intercept illegal information on large, country-wide private computer networks, "such as those set up by terrorist groups like the IRA [Irish Republican Army]". But, he says, "they allow the intelligence agencies access to confidential medical records that have no relevance to crime".

Under the new US proposals, the government has agreed to temporarily lift the ban on the export of stronger — but non-escrowed — encryption products that use a key length of 40 bits. (The strength of a product is measured by the length of its key.) But in return, software manufacturers will be asked to drop their opposition to key escrow and incorporate it into future encryption products.

Software companies have been concerned that the government's key escrow proposals would jeopardize their US\$600 million revenue from exports of encryption software. Manufacturers pointed out at the time that international customers were unlikely to buy products that would entitle the US government to eavesdrop on them.

Vice President Al Gore said the administration will not be supporting a Republican bill that calls for the lifting of export controls on all encryption software. "We do not want to do anything that would damage our own national security or public safety by spreading unbreakable encryption."

Ehsan Masood

High standards urged for French charities

Paris. French researchers were warned last week to observe the highest standards of openness and scientific integrity in their dealings with research charities, and to ensure that such charities are run with exemplary levels of transparency and quality.

The warning was issued in a report by an ethics committee of the National Centre for Scientific Research (CNRS), set up in 1994 to identify ethical problems in research. The 14 members of the committee, known as COMETS, are drawn from academic institutions, the legal profession and other sectors. They include the former science minister and CNRS director Hubert Curien, and the panel is chaired by Hélène Ahrweiler, a former rector of the Academy of Paris.

The committee's report was prompted by a recent financial scandal at France's largest medical charity, L'Association pour la Recherche contre le Cancer (ARC) (see *Nature* 379, 103; 1996). Its 14 recommendations, proposed as a code of conduct for researchers working with charities, represent a belated attempt to close some of the glaring gaps in the regulation of French charities

that were exposed by the ARC scandal. "No cause, however sacred, authorizes the sacrifice of moral values," says the report, an implicit criticism of the fact that many scientists associated with ARC had long turned a blind eye to activities at the charity, even defending it against allegations of mismanagement and wrongdoing.

The ethics committee recommends that scientists who sit on the boards of charities should avoid acting as both "judge and jury", and that funds should be distributed by independent scientific committees. (The executive board of ARC included many members who received large amounts of funding from the charity.)

Members of such committees should ensure that charities distribute grants fairly, and evaluate how money is spent. If researchers suspect wrongdoing, they should "not hesitate" to withdraw their support for a charity and to "make public their disagreement", says the report, in an explicit call for more whistleblowing.

Researchers themselves should submit funding received from charities to the rigor-

ous accounting procedures used by the research organizations to administer their budgets, it adds. This is already a legal obligation. But many researchers have operated separate accounts to administer funds they receive from charities; an audit of ARC found that many recipients of the charity's grants were unable to provide receipts to confirm how they had spent the money.

The report also warns researchers against colluding with aspects of fundraising that could mislead the public and so damage the image of research. Scientists should be careful not to raise unduly the public's hopes of medical progress; they should rather explain the long-term nature of research and the need to support fundamental research.

The thrust of the report is that, although funding of research by charities is increasingly important in France, it is only a fraction of the total state investment in science. Researchers need to avoid becoming dependent on outside associations, and should ensure that they do not unduly influence the direction of their research or policies, says the report. **Declan Butler**

Sceptics course seeks to stem 'tide of irrationalism'

Boston. Eager to train the next generation of qualified sceptics, the Center for Inquiry Institute — a branch of the Committee for the Scientific Investigation of Claims of the Paranormal (CSICOP) — has launched a three-year academic programme leading to a certificate of proficiency in "science and the paranormal" and "humanistic studies".

The aim is to boost the ranks of those prepared to challenge what some describe as a "rising tide of irrationalism", with 48 per cent of Americans believing in unidentified flying objects, according to a recent poll, and 29 per cent convinced that contact has been made with aliens.

The academic programme was announced last month at the First World Sceptics Congress, held at the State University of New York at Buffalo, in celebration of CSICOP's twentieth anniversary. A Council for Media Integrity on Science, aimed at countering distorted views of science presented in the media, is also being set up.

The Center for Inquiry Institute has been running workshops, seminars and conferences since it was formed in 1987. But Paul Kurtz, the chairman of CSICOP, says that the more formal programme will be the "first effort to develop the tools of sceptical thinking in a systematic way and to apply it to a wide range of questions".

To complete the programme, students will need to acquire 30 credits through courses, workshops and a research paper or thesis. "How fast people can graduate depends on how fast we can offer courses,"

explains Joe Nickell, associate dean and a senior research fellow at CSICOP.

The centre is exploring the possibility of gaining accreditation for the programme, and is an accredited research institute of Westminster College, associated with the University of Oxford, Nickell says. "My feeling is that not many sceptics are planning to take our courses for college credit, as so many of them are beyond college."

Eugenie Scott, a physical anthropologist and executive director of the National Center for Science Education in Berkeley, California, thinks the programme will be worthwhile, regardless of credits or accreditation. "Studying paranormal claims is a good way to learn critical thinking and to learn about the scientific method — the need for empirical evidence, experimental

controls and so forth," she says. "Studies have shown that individuals who understand science live in a less fearful world."

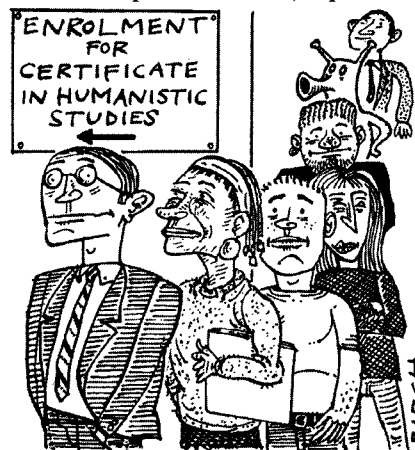
But some sceptics are themselves somewhat sceptical of the initiative. Ray Hyman, for example, a psychologist at the University of Oregon, and a well-known debunker of extra-sensory perception and other paranormal phenomena, says it unlikely that his successor will emerge from the programme.

"You need a lot of training in statistics and laboratory work to do what I do, and you're not going to get that by taking a few courses," he says. "While I have my doubts as to why people need formal certificates, I do think that sceptics need training." One problem is that "there are no ground rules or manuals on how to be a good sceptic". Hyman believes that critical thinking can be taught. But it is even more important to learn how to get good information.

Communication skills are crucial in this field, Hyman maintains, and courses on the subject are being offered by the centre. "We've had more trouble with people who claim to be sceptics talking out of turn, speaking without knowing what they're talking about, than we've had with paranormal believers," he says.

Manners are important too, he stresses. "The trick is how to tell people they're wrong without putting them down. Sceptics have to be civilized in their delivery, because we won't reach the 'believers' by attacking them and making them more defensive."

Steve Nadis



House rejects bid to lift ban on funding for embryo research

Washington. In a defeat for liberal Democrats, the US House of Representatives last week blocked an attempt by Nita Lowey (Democrat, New York) to remove from a 1997 funding bill — which would boost the budget of the National Institutes of Health by 6.9 per cent, to \$12.7 billion — a blanket ban on federal funding for human embryo research.

Lowey had initially succeeded in removing the ban in a House appropriations subcommittee (see *Nature* **381**, 633; 1996). But the full committee reversed the decision. She tried again on the House floor, introducing an amendment lifting the ban except in the case of embryos specifically created for research purposes. But the House voted by 256 to 167 against the amendment. The spending bill itself was narrowly approved, and now awaits action in the Senate. □

Peers criticize Gulf War choice

London. Members of Britain's House of Lords have criticized a decision by the US government to fund a research project into Gulf War illness at King's College Medical School in London. During a debate last week, The Countess of Mar, patron of the Gulf Veterans Association, said the US Department of Defense (DOD) ought not to have awarded the funds on the grounds that the main researcher, Simon Wessley, held "preconceived ideas" that Gulf War syndrome "is caused solely by post-traumatic stress syndrome".

Earl Howe, parliamentary under-secretary of state for defence, said the government had been in "close touch" with the US authorities, and had no reason to doubt the DOD's decision. Wessley's project is one of 12 selected for funding by the research working group of the US Persian Gulf Veterans Coordinating Board. □

German biotech strategy falters

Munich. A recent boast by Jürgen Rüttgers, Germany's research minister, that his country would become the leading force in biotechnology in Europe by the turn of the century, is already starting to look shaky. This goal had assumed an increase in public funding of five to six per cent a year, accompanied by a much sharper increase in industrial financing.

The 1997 budget, however, sees only a three per cent increase for biotechnology research (see page 197), and industry appears unlikely to fill the gap. Nevertheless a spokesman for the research ministry's biotechnology unit remains optimistic. "Given the situation, we are pleased that there is any increase at all," he says. □

US cancer council planned

Washington. The governing board of the US National Research Council (NRC) is expected to agree this week to establish a new National Cancer Council to help coordinate cancer research and advise the government and the public on controversial issues surrounding cancer treatment. Richard Klausner, the director of the National Cancer Institute, says that the council will be "as free from institutional biases as possible", so as to inspire public confidence. It is expected to be set up as a permanent body within the NRC, the operating arm of the National Academy of Sciences. □

Chirac backs epidemiology

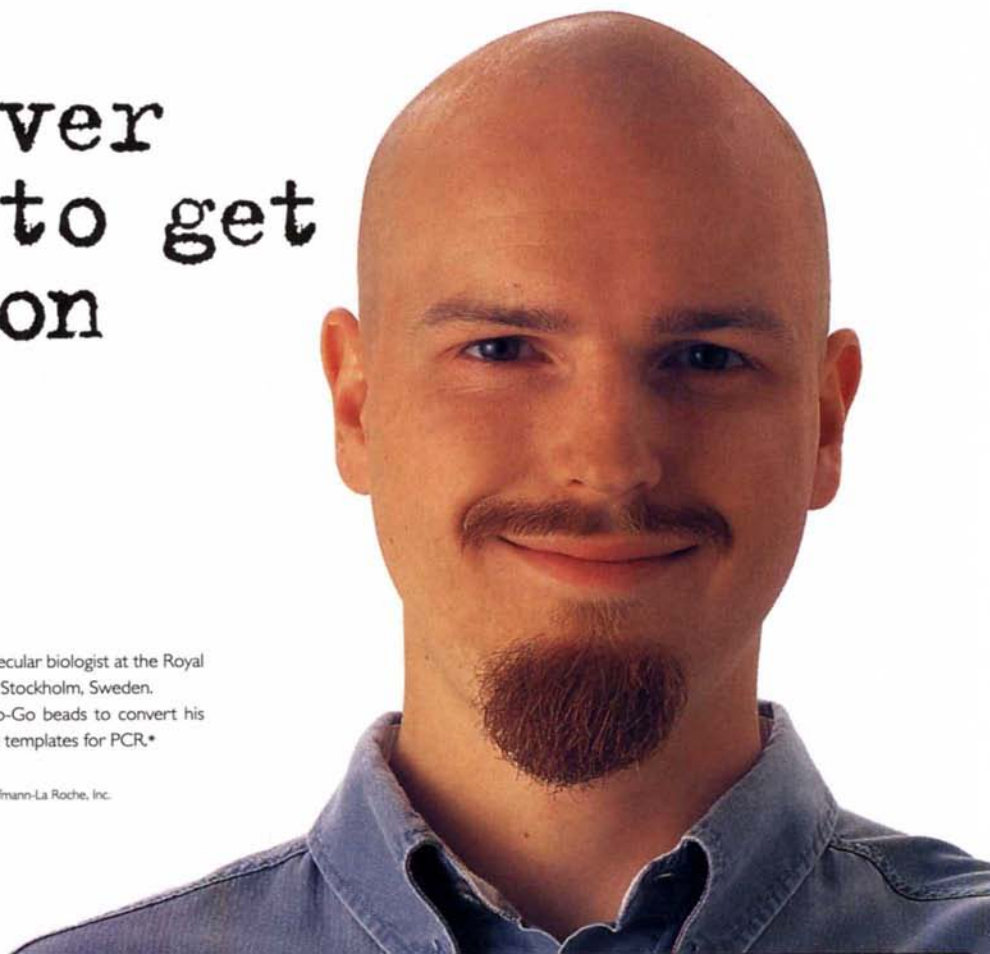
Paris. France must make a major effort to tackle the threat of emerging and infectious diseases, according to Jacques Chirac, the French president. Speaking in a televised interview on Bastille Day (14 July), Chirac deplored what he described as the current excessive emphasis on molecular biology, and said he would instruct the government of Alain Juppé to ensure that microbiology and

Patrik never
fails to get
a reaction

Patrik Samuelson is a molecular biologist at the Royal Institute of Technology in Stockholm, Sweden.

Patrik uses Ready-To-Go beads to convert his RNA samples into cDNA templates for PCR.*

* PCR is a patented process of Hoffmann-La Roche, Inc.



epidemiology appeared higher on the research agenda.

How far Chirac's wish will be translated by the government into action remains to be seen. But it seems clear that Chirac's comments continue the recent effort by the administration, in the wake of the crisis over bovine spongiform encephalopathy, to be seen to be closely tuned to scientific developments affecting public health (see *Nature* 382, 5; 1996). For his part, Chirac confided on Sunday that he and his family continue to eat beef "in all tranquility". □

India clears gene therapy project

New Delhi. The Indian government has given the go-ahead for the country's first research project on gene therapy. A four-member team led by Rita Mulherkar, a French-trained microbiologist, will attempt to develop a 'suicide' gene therapy to treat oral cancer, the most common cancer in India. Four more gene therapy projects are in the pipeline.

Manju Sharma, secretary to the Department of Biotechnology (DBT), which is funding the research at the Tata Memorial Centre in Bombay, says that many proposals have been received. "But none could be funded because the ethical guidelines were not ready until early this year." Now that hurdle has been removed, DBT expects a boom in gene therapy research. □

British Biotech shares slump

London. The stock market value of British Biotech, the small biotechnology company whose value rocketed to £37 (US\$56) per share last month after its development of a cancer treatment, went into free-fall last week. Its shares were valued at below £21 on Monday, and continued to fall up to the close of trading on the London Stock Exchange. A planned £143-million rights issue has been described by one city commentator as being "perched on the edge of disaster". The news has also hit other biotechnology shares. □

Australian institute restructures

Sydney. Having scrapped its previous separation into six institutes, Australia's national research agency, the Commonwealth Scientific and Industrial Research Organization (CSIRO), is to be further restructured by mergers among its 33 specialist research divisions and two centres. The new chief executive, Malcolm McIntosh, justifies the move on grounds of "focusing scientific resources in areas of national priority", and diverting more funds into supporting economic competitiveness and studies of the environment and oceans. A new category of research staff, known as Corporate Fellows, will lead large scientific teams without heavy administrative loads. □

Industry costs climate moves

London. The fossil fuel and energy lobby has stepped up its campaign to persuade governments to think twice before reducing greenhouse gas emissions, faced with continued protests from



Climate Action Network

environmentalists about the likely impacts of global warming — including sea-level rise (see picture above).

Computer models commissioned by the Global Climate Coalition and the American Petroleum Institute, and on show at the current intergovernmental climate talks in Geneva, forecast economic decline in both developed and developing countries if industrialized countries commit themselves to a significant decline in carbon dioxide emissions. But the models have been criticized as "deeply flawed" by the Climate Action Network, an alliance of environmentalist organizations. □

Do you always get a reaction when converting RNA into single-stranded cDNA templates for RT-PCR? Patrik does—thanks to a revolutionary polymeric bead that makes these critical conversions for him in new Ready-To-Go™ Kits.

The kits utilize the new Ready-To-Go bead that avoids many potential errors common to handling aqueous reagents.

That's because the "bead" is a complete,

pre-formulated, single-dose reaction—so there's little chance of your PCR templates becoming contaminated.

Just add your sample to the ambient-stable bead and watch it dissolve—that's all the pipetting required. In under a minute, you can begin full-length first-strand cDNA synthesis. This new approach to producing cDNA is only available from Pharmacia Biotech.

Just call us at 1 (800) 526 3593 in the United States, or +46 18 16 5011 from the rest of the world, for more information. Ask about the bead that never fails to give you a reaction.

 **Pharmacia
Biotech**
Uppsala, Sweden. (And the rest of the world)

Something up Dawson's sleeve?

SIR — I should like to add to Henry Gee's summary of Brian Gardiner's suspicions about Martin Hinton as the forger of the Piltdown fossils¹. Just a few years ago, Frank Spencer² accused Sir Arthur Keith of being the forger. Anyone who knew Keith well would not entertain such a ridiculous verdict, but Spencer had him guilty from the first sentence, and then, with winks and nods, pinned him down as the forger. But Spencer could have incriminated any of the suspects he mentioned, including Hinton, for the book reads like a 'whodunit'. Now Hinton is accused of being the forger because he invented the stain found on the so-called fossils supposedly recovered at Piltdown.

There are three questions to be answered. When did Hinton invent the stain? Who else could have had access to the stain? Was Hinton knowledgeable enough in human anatomy to so damage the jawbone in places vital for distinguishing between human and simian bones?

If Hinton stained the bones, he did so when he was a volunteer worker at the Natural History Museum. But the stained bones had to be transported to Piltdown — or had they? What is the proof that the bones picked up at Piltdown were the forged fossils demonstrated at the museum?

What has been established is this: Arthur Smith Woodward, the keeper of palaeontology at the museum, and Charles Dawson were present at Piltdown when the cranial fragments were found; and Smith Woodward and Dawson were present at Piltdown when the right half of the mandible was found.

There are reports by Smith Woodward and Dawson testifying to these statements. According to what Smith Woodward wrote many years later, and introduced in Spencer's book² (page 34): "Mr Dawson was exploring some untouched remnants of the original gravel at the bottom of the pit, when we both saw the half of a human lower jaw fly out in front of the pick-shaped end of the hammer which he was using."

Now either that statement is a true account of the discovery of the mandible or it is not. If it is true and Hinton is the supposed forger, then he must have had access to the pit and placed (in "untouched remnants") the mandible at the bottom of the pit some four feet down. Hinton was one of the Ightham Circle men and so could have bribed the farm worker Venus Hargreaves to plant the mandible and consolidate the soil afterwards. All that takes some believing. On the other hand, Dawson could have concealed the mandible in the sleeve of his shirt.

Spencer was unable to find any evidence

that Hinton had any direct dealings with Dawson. Keith points out in his autobiography³ (page 324): "As a palaeontologist Smith Woodward enjoyed, and deserved, the highest reputation, but he had no special knowledge of the human body". If we consider what Keith stated about the mandible in his *Antiquity of Man*⁴ we find that he was sceptical from the moment he was shown the bones.

Here are a few of his comments: "In the Piltdown mandible the conformation is that of an ape" (page 507); "Piltdown jaw should be linked with a skull which is distinctly more anthropoid than the one actually found"; "The very part of the jaw — the condyle — which we most need to give us the form of contact with the skull is broken away" (pages 645–647); "We have seen that in many features the Piltdown mandible resembles that of the chimpanzee" (page 652).

It may be that the mandible really was broken at the vital distinguishing parts when it was admitted to the museum collection some time previously. Someone like Smith Woodward would not be aware of the significance of the missing parts, but a skilled human anatomist, such as Keith, would, as he pointed out on many occasions.

Dawson carried out staining techniques as indicated by the contents found in Harry Morris's cabinet. The Piltdown bones were found by Dawson and Smith Woodward and presented by them to the scientific establishment. Need we look any further for the forgers?

W. J. Dempster

Wavertree, Tytherley Road,

Lockerley, Hampshire SO51 0LW, UK

1. Gee, H. *Nature* **381**, 261–262 (1996).
2. Spencer, F. *Piltdown — A Scientific Forgery* (Natural History Museum/Oxford Univ. Press, 1990).
3. Keith, A. *An Autobiography* (Watson, London, 1950).
4. Keith, A. *Antiquity of Man* (Williams & Norgate, 1929 edn).

Is the Pope an alien?

SIR — A syllogism is a form of deductive reasoning which, contrary to the claim of Beck-Bornholdt and Dubben¹, is not the model for the inductive reasoning of a test of significance.

Tests of significance are usually reasonable because an improbable event alerts us to the possibility that the hypothesis on which the calculation of the probability was based might be able to be replaced by a better hypothesis leading to a higher probability. Obviously, the greater the original improbability (*P* value small) the

greater the scope for improvement.

In the case of Beck-Bornholdt and Dubben's bizarre example, the conclusion is not that Pope John Paul II is not a human being but that the null hypothesis that popes are selected at random from the world population is false.

It is not correct to state that "no alternatives are available" to tests of significance. The likelihood approach to statistical inference, which encapsulates the idea of assessing rival hypotheses according to their probabilities of generating the data, is not only a non-Bayesian alternative to the 'repeated sampling' methods of which tests of significance are an example, but is itself the best explanation for why tests of significance are usually reasonable procedures.

Likelihood was first given serious consideration by R. A. Fisher² (indeed, he gave the word its technical meaning long before he coined the phrase 'null hypothesis'), and has increasingly come to play an important role in science, especially in the biological sciences. I discussed the relationship between likelihood and tests of significance many years ago in Chapter 9 of *Likelihood*³.

A. W. F. Edwards

Gonville and Caius College,

Cambridge CB2 1TA, UK

1. Beck-Bornholdt, H.-P. & Dubben, H.-H. *Nature* **381**, 730 (1996).
2. Fisher, R. A. *Statistical Methods and Scientific Inference* (Oliver & Boyd, Edinburgh, 1956).
3. Edwards, A. W. F. *Likelihood* (Cambridge Univ. Press, Cambridge, 1972; Johns Hopkins Univ. Press, Baltimore, 1992).

Right to publish

SIR — You report in a recent News story (*Nature* **381**, 458; 1996) the controversy surrounding the publication of a scientific paper from a jailed professor.

Jailing a criminal is a punishment and serves to prevent the criminal from committing his crime again. It is also hoped that imprisonment may serve to transform and educate the criminal into becoming a useful and honest member of society even from within the prison.

If a jailed criminal produces a work useful to society, should we not permit society to use this work? If a criminal writes a sound scientific communication, should not we allow it to be published?

In my opinion, publication should be permitted provided that the criminal does not profit financially.

Zoltan Ovary

Department of Pathology,

NYU Medical Center, 550 First Avenue,

New York, New York 10016, USA

corres@nature.com

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to corres@nature.com

A revolutionary philosopher of science

David L. Hull

Thomas S. Kuhn, who died on 17 June aged 73, is best known for introducing the principles of 'paradigms' and 'incommensurability' into the study of science. The influence of his work, however, goes beyond such slogans.

TOGETHER with Karl Popper, who died in 1994, Thomas Kuhn has been the most influential philosopher of science in recent history. His main opus, *The Structure of Scientific Revolutions* (1962), has been translated into 25 languages, and the English edition alone sold over a million copies.

Just as Popper is identified with the principle of falsifiability, the notion of a 'paradigm' brings Kuhn to mind. In *Structure*, Kuhn distinguished between normal and revolutionary science. In periods of normal science the basic outline of the world view (or paradigm) held by a group of scientists is fairly well understood. Scientists know what they are supposed to do and how they are to do it. For example, Mendelian geneticists establish pure lines, cross them, cross the results and see what sorts of ratios they get. Later Mendelian geneticists did not have to be geniuses to make significant contributions to Mendelian genetics. All they had to do was follow the exemplar set out by their forerunners.

Revolutionary science is quite different because scientists are questioning the very ground upon which they are standing. In fact, a period in science is termed 'revolutionary' to the extent that the basis of that science, possibly science as such, is being questioned. During scientific revolutions, scientists disagree not only about matters of fact (for example whether species evolve or not) but also about science itself. Can genuine science make reference to God and miracles, or must all such references be excised from scientific publications? Must scientists assume that the natural forces that are currently at work were the only forces that functioned in the past? Much more trivially, must scientists always write in the passive voice?

In the past, those of us interested in science have tended to emphasize periods of revolutionary change—the radical changes stemming from the concepts of continental drift and plate tectonics, for instance, or the transition from Aristotelian to Newtonian physics, and from there to the notions of relativity and quantum theories. But one reason why many scientists liked Kuhn's work was that they recognized his idea of normal science in their own experience. Most of the time,

most scientists are engaged in normal science. Anyone attempting to explain science had better pay attention to the typical case, instead of concentrating totally on the rare, abnormal periods of revolutionary change.

Scientists also identified with Kuhn because he urged a greater role for real examples in evaluating alternative views of science, in particular examples drawn from



the history of science. Indeed, it was reading some history of science for the first time that radicalized Kuhn. The little history that he knew as a young physicist was the sort of whiggish history that makes all early science zero in on the present. When he read a little real history, he discovered that it was not at all what he had been led to expect from his training as a scientist.

Kuhn himself began by studying the Copernican revolution, in which the Sun displaced the Earth as the centre of the Solar System. One of the problems he confronted is strikingly exemplified in the preceding statement. Quite obviously, the geocentrists did not think that the Earth was at the centre of the *Solar* System. That we now term the system formed by the Sun, Moon, Earth and other planets as such indicates that we think that the Copernicans were right. That is why we have adopted their ter-

minology, and not that of the geocentrists.

After publishing *Structure*, Kuhn devoted much of his energy to working on the history of physics, in particular the early development of the quantum theory. Surprisingly, his own general views about science played only the most minimal role in his later publications. Other historians, for their part, have tried to apply Kuhn's views in their own work. Can the history of science be interpreted as an alternation between normal and revolutionary science? Were the introduction of Copernican theory, cell theory or evolutionary theory instances of 'revolutions' in Kuhn's sense? The usual conclusion that historians have been forced to reach is that Kuhn did not articulate his views about scientific revolutions with sufficient precision to allow them to decide.

However, the slogan with which Kuhn is most closely identified is that alternative paradigms are incommensurable. Although 'incommensurable' has not become a cocktail party buzzword, 'paradigm' certainly has, so much so that editors of more serious journals tend to delete any reference to paradigms from manuscripts. According to Kuhn, all terms, even the most observational, are theory laden. Even though two different theories include references to 'mass', this theoretical notion may be defined differently in the two theories. As a result, some of these differences in meaning are incorporated even in apparently straightforward descriptions of observations. As a simple example, Anton van Leeuwenhoek prided himself in describing what he saw through his lenses, not in theorizing about them. Yet he termed the little creatures that he saw 'animalcules', implying what sorts of creatures he thought they were—little animals, not little plants.

One consequence of the theoretical differences between different paradigms, and the notion that even observational statements are theory laden, is that two theories cannot be compared with absolute precision by reference to data. In addition, the chasms that characterize some revolutions are so deep that even the notions of 'data' and 'inference' are themselves at issue. For example, Freudians accept reports of introspection as data, but behaviourists

certainly do not. Some scientists insist that only deductive inferences are acceptable in science. Others are willing to reason probabilistically.

As a result of all of the above, scientific paradigms are incommensurable: scientists cannot decide between alternative paradigms solely on the basis of reason, argument and evidence. Moreover, all sorts of other factors play a role in theory change in science. Kuhn even goes so far as to compare scientific revolutions to political revolutions, and no one thinks that the latter are entirely a matter of reason, argument and evidence.

Professional philosophers of science were put off by Kuhn's views, especially his principle of incommensurability. Philosophers value conceptual and terminological precision, and early critics found Kuhn's style too poetical for their tastes. For example, Kuhn alternated between saying that different paradigms are incommensurable and incompatible, but if paradigms are genuinely incommensurable, then they cannot, strictly speaking, contradict each other. That is the whole point of incommensurability. Incommensurability even brings into question the issue of which theories count as 'alternatives' to each other. Kuhn was deeply frustrated by the philosophical responses to his views — so much so that he claimed that, among all the readers of his work, philosophers were uniquely unable to understand him.

Scientists, for their part, have tended not to go into any great details about incommensurability, except to use it to deflect criticisms raised by their opponents. "You don't understand because you are too enmeshed in the old paradigm. Once you make the requisite leap of faith, all will be clear." In some societies genital mutilation of young girls is perfectly acceptable. In other societies it is not. No outside, objective perspective exists for evaluating such practices. Any attempt by one society to impose its mores on another is just cultural imperialism. Similarly, scientists attempting to impose a scientific world view on the rest of us is just one more instance of cultural imperialism. Many of Kuhn's readers interpret him as advocating a similar relativism with respect to knowledge.

Social relativism has a long history. At the moment it has become popular again, in part fuelled by an increasing hostility to science. As scientists have become more powerful, people from both the religious right and the political left have raised emotional protests as they feel themselves being marginalized. Evolutionary theory is 'only' a theory and should be taught along-



side other theories, in particular Creationism. Bible stories should be taught in biology classes. Similarly, sociobiology is not to be evaluated entirely in terms of reason, argument and evidence. Its social implications and implicit social assumptions must also be weighed in the balance. In certain situations, 'good' might take precedence over 'true'. In fact, there is no such thing as truth. Science is nothing more than people who have different interests fighting it out. Like it or not, might determines right.

In this new battle within science studies, those who view science as relative to time and place have adopted Kuhn as one of their patron saints. Because the transition from one paradigm to another cannot be explained entirely in terms of reason, argument and evidence, such factors play no role whatsoever in such transitions. Evidence is infinitely pliable, little more than silly putty to be moulded as one sees fit —

"If Kuhn is interpreted with only a modicum of charity, he has had quite a bit to tell us about the nature of science."

or so runs the argument. Kuhn himself was dismayed by his relativist disciples. This was not what he intended at all!

One of the refrains in the social studies of knowledge is reflexivity. If something is true of everyone, then it is true of the person making this claim. If evidence plays such a minimal role in scientists making the decisions that they do, then it can play no larger a role in the decisions made by those of us who study science. But why then do social constructivists spend so much time gathering evidence to show how irrelevant evidence is, including the evidence provided by case studies in the sociology of knowledge?

The all-purpose whipping boys in the science studies literature are the positivists. These are evil people who claim that evidence plays a major role in science. Strangely enough, however, they do not think that it plays all that much of a role in their own deliberations about science. Conversely, those social relativists who devote so much of their time to case studies are put in the equally problematic position of claiming that evidence does not play much of a role in science but is decisive in the study of science. To be consistent, social relativists must view appeals to evidence both in science and in the study of science as rhetorical strategies designed to take advantage of the biases of the time.

Only one further alternative exists — evidence just might play a significant role both in science and in the study of science. On one reading of Kuhn, this is exactly what he proposed in calling for the larger involvement of the history of science in the philosophy of science. It also explains all

the energy that he invested in studying science. As real as all the problems that Kuhn raises about testing general positions against evidence actually are, they can be overcome — both in science and in the study of science.

As social relativists have become increasingly vocal in their opposition to what they take to be simplistic, positivist views of science, more traditional philosophers of science have come to conclude that their early, highly critical reaction to Kuhn was too one-sided. If Kuhn is interpreted with only a modicum of charity, he has had quite a bit to tell us about the nature of science. The main reason for members of professional societies electing one of their own as president is to show that they value the contributions that this expert has made. The fact that many younger, less influential philosophers (including myself) were elected president of the Philosophy of Science Association before Kuhn was elected in 1988, indicates both the early reservations that philosophers had about Kuhn's philosophy and the change of heart that had occurred. (I still remember Kuhn taking me to one side after he was installed as president, to ask my advice about which person would be right for which committee. He really did not know very much about others in the association except for their publications.)

During this period Kuhn was also busy re-evaluating his own contributions to the philosophy of science, working on a comprehensive synthesis of his views that would resolve the questions that traditional philosophers have raised to them and deflect the relativist misinterpretations. Science turns out to be very different from what Kuhn had been led to expect on the basis of his undergraduate training in physics, but it is not just another weapon for those in power to subjugate the disenfranchised masses either. This book was essentially complete at the time of Kuhn's death, and plans are under way for its publication by the University of Chicago Press.

What I find amazing about Thomas Kuhn is that one person could have influenced so many different sorts of people in so many different ways. I suspect that a hundred years from now, Kuhn will be one of the few philosophers of science who will be looked back upon as having radically changed our understanding of science. Who knows what later scholars will take that change to be? Too many culturally significant figures are destined to be remembered for overly simple slogans, such as Mies van der Rohe's "less is more". It would be a great shame if Kuhn were remembered only for his claim that different paradigms are incommensurable with each other. □

David L. Hull is in the Department of Philosophy, Northwestern University, Evanston, Illinois 60208, USA.

Catching the inner core in a spin

Kathy Whaler and Richard Holme

The inner core of the Earth, a solid piece of iron 2,400 kilometres in diameter, is rotating faster than the rest of the planet. At its equator, it moves under our feet by more than 20 kilometres a year.

SEISMIC waves generated by earthquakes and nuclear explosions travel through the Earth at different speeds, depending on the material they pass through and, in some cases, on the direction of travel. By carefully analysing thirty years of seismic data, Song and Richards show on page 221 of this issue¹ that the inner core rotates relative to the mantle and crust by about 1.1 degrees per year. This is not just a grand curiosity; it has implications for the dynamics of the whole planet and for the generation of the Earth's magnetic field.

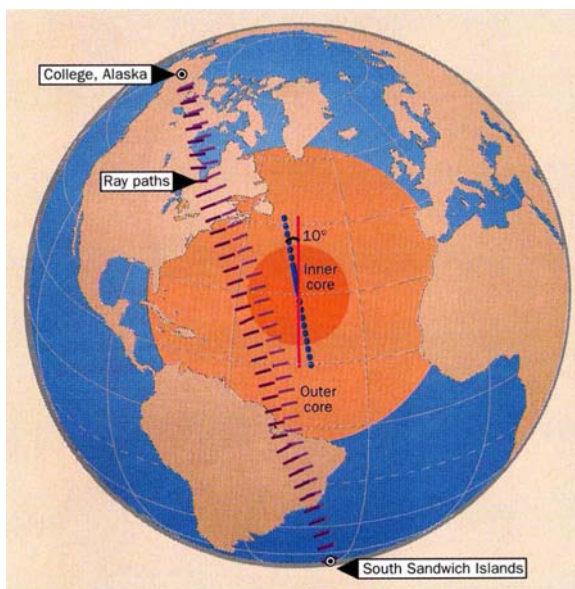
It is well established that the inner core is anisotropic — the time taken for seismic energy to propagate through it along a north-south path is less than that for a path in the equatorial plane. More recently, evidence has grown that the axis of anisotropy (the 'fast' axis) is at an angle of about 10° to the north-south axis of the Earth. These results are based on studies of the difference in travel time between a wave that propagates through the inner core and one that takes a slightly different path, going only as deep as the outer core. In this way, seismologists can rule out irregularities in the mantle and local structure around the sources and receivers of the seismic energy as the cause of their observations, because the seismic waves follow nearly the same path through most of the Earth. The cause of the anisotropy is less certain, but the suggestion is that the crystalline structure of the inner core has a preferred axis².

Song and Richards find systematic differences over the years in the differential travel times of waves that follow a roughly north-south path through the core. For example, the differential travel-time for earthquakes in the South Sandwich Islands recorded at College, Alaska, was about 0.4 seconds less at the start of the records they studied, in 1967, than for an earthquake in 1995. In contrast, there are no discernible changes for waves travelling near to the equatorial plane, such as those from earthquakes in the Tonga region recorded at the Graefenberg array in Germany.

These observations can be explained by

the symmetry axis of the anisotropy of the inner core rotating by about 1.1° per year in an easterly direction relative to the mean rotation of the Earth — the axis of symmetry intersects the Earth's surface at 79° N, 169° E today, whereas thirty years ago it was 33° farther west.

The solid inner core of the Earth grows



Comparative seismology: two seismic waves from an earthquake in the South Pacific travel to Alaska by slightly different paths. Over the years, the timed results of this race have changed, betraying the inner core's rotation.

slowly with time as the Earth cools. It freezes out almost pure iron from the outer core liquid³, which is about 90% iron and 10% lighter elements — exactly which lighter elements has been the subject of a great deal of controversy over the years. The residual liquid is enriched in the lighter elements and heated by the latent heat of freezing iron, and so is more

buoyant than the rest of the outer core. Its rise drives convective motions that generate the magnetic field by self-sustaining dynamo action⁴ — the magnetic field lines get wound up by the motion, so kinetic energy is converted into magnetic energy. Hence, the formation of the inner core provides the energy source for the Earth's magnetic field. Numerical studies have shown the importance of the inner core to the geodynamo⁵, and recent dynamical, three-dimensional simulations^{6,7} produce an inner core rotating relative to the outer core at a similar rate to that deduced by Song and Richards, also in an easterly direction. This rotation rate is of the same order of magnitude as convective flow speeds.

The magnetic field threads the inner core, coupling inner and outer cores electromagnetically. They are also more weakly⁸ coupled viscously, gravitationally and by pressure differences exerted by the flow, and perhaps also topographically, if the surface of the inner core is not perfectly smooth. If the rotation rate is constant, there must be no net torque on the inner core. That might simply reflect a balance between electromagnetic torques, as considered in a simplified model of differential rotation of the inner core⁹. Alternatively, although it is usually assumed that the outer core viscosity is too small for viscous torques to be important, a postulated 'mushy zone' at the base of the outer core¹⁰ could have a higher viscosity.

The mantle, however, does not rotate at a constant rate, as we can see from changes in the length of day¹¹. Torques

1. Song, X. & Richards, P. G. *Nature* **382**, 221–224 (1996).
2. Jeanloz, R. & Wenk, H. R. *Geophys. Res. Lett.* **15**, 72–75 (1988).
3. Masters, G. *Geophys. J. R. astr. Soc.* **57**, 507–534 (1979).
4. Gubbins, D., Masters, T. G. & Jacobs, J. A. *Geophys. J. R. astr. Soc.* **59**, 57–99 (1979).
5. Hollerbach, R. & Jones, C. A. *Nature* **365**, 541–543 (1993).
6. Glatzmaier, G. A. & Roberts, P. H. *Phys. Earth planet. Inter.* **91**, 63–75 (1995).
7. Glatzmaier, G. A. & Roberts, P. H. *Nature* **377**, 203–209 (1995).
8. Szeto, A. M. K. in *Dynamics of the Earth's Deep Interior and Earth Rotation* (eds Le Mouél, J.-L., Smylie, D. E. & Herring, T.) 31–33 (AGU, Washington DC, 1993).

9. Gubbins, D. *J. geophys. Res.* **86**, 11695–11699 (1981).
10. Fearn, D. R., Loper, D. E. & Roberts, P. H. *Nature* **292**, 232–233 (1981).
11. Morrison, I. V. *Geophys. J. R. astr. Soc.* **58**, 349–360 (1979).
12. Jault, D., Gire, C. & Le Mouél, J.-L. *Nature* **33**, 353–356 (1988).
13. Mathews, P. M., Buffett, B. A., Herring, T. A. & Shapiro, I. I. *J. geophys. Res.* **96**, 8243–8257 (1991).
14. Bloxham, J., Gubbins, D. & Jackson, A. *Phil. Trans. R. Soc. A* **329**, 415–502 (1989).
15. Halley, E. *Phil. Trans. R. Soc.* **13**, 208–221 (1683).
16. Bullard, E. C., Freeman, C., Gellman, H. & Nixon, J. *Phil. Trans. R. Soc. A* **243**, 61–92 (1950).

between mantle and atmosphere, and mantle and core, must conserve the total angular momentum of the Earth-atmosphere system, so that if the mantle slows down, the atmosphere and/or core must speed up. Exchanges of angular momentum with the atmosphere take place on a timescale of days to years, whereas the core accommodates the observed changes in day length that are on a timescale of decades. In one proposed mechanism¹², the core responds to these changes by axis-symmetric flow around cylinders centred on the rotation axis, the flow being the same at all points a given distance from the axis. A rotating inner core could disrupt this simple situation, as the cylinder tangential to it would be affected by their relative motion at the equator of a few kilometres per year. In previous studies, angular momentum changes have been considered to be between the mantle and core as a whole, with the inner core a passive partner, but that would require a rethink if a more detailed study showed that the inner core rotation rate varied too.

If the symmetry axis of the inner core were aligned with the axis of anisotropy, its angular momentum would not lie along the axis of rotation, but would rotate with it. That ought to have an observable effect on the Earth's rotation¹³, but very-long-baseline interferometry has failed to detect any effect on the required timescale, and so the symmetry axis is probably close to the axis of rotation after all.

Even so, the dynamical consequences of Song and Richards' result could be important. Maps of the changing magnetic field at the core-mantle boundary¹⁴, which trace the fluid velocity, suggest that westward flow at the surface of the outer core is responsible for the westward drift of the magnetic field first observed by Halley¹⁵, even though the flow is confined to a limited area of the core surface. If changes in angular momentum of the core average to zero over long timescales, westward drift could be explained by the outer part of the core rotating less rapidly than the inner part¹⁶: westward motion at the core-mantle boundary would imply eastward motion at depth, which is what Song and Richards observe for the inner core.

Digital seismological data are more accurate and much easier to work with than are the older analogue records. But now we know that we cannot throw away all those old paper records, because the Earth's seismological structure is changing. We need the older analogue records to confirm, and determine more accurately, the relative rotation of the inner core. □

Kathy Whaler and Richard Holme are in the Department of Geology and Geophysics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK.

... harmful to the brain

Leslie L. Iversen

THE view that nicotine is a powerful addictive drug was summarized in the influential US Surgeon General's Report in 1988¹. This assessment has been strengthened by further research showing similarities between nicotine and a number of drugs of addiction including cocaine, amphetamine and heroin². But it was not always thus, and there remains the view that nicotine — and thus smoking — is not addictive, with all that implies for anti-smoking campaigns.

A paper elsewhere in this issue (page 255³), by Di Chiara and colleagues at the University of Cagliari, adds new weight to

intact brains of conscious animals show a selective release of dopamine in the nucleus accumbens in response to administration of psychostimulants and morphine. Furthermore, damage to the dopamine fibres in the nucleus accumbens in animals trained to self-administer addictive drugs leads to extinction of this behaviour.

It has been known for some time that nicotine can mimic the addictive psychostimulants in causing dopamine release from the nucleus accumbens⁵, as shown by the Cagliari group more than ten years ago⁶. The new results add a further

Dor Miller/Network

refinement, in that separate measurements of dopamine release and energy metabolism (measured by accumulation of radiolabelled deoxyglucose) were made in the subdivisions of the nucleus accumbens known as the 'shell' and the 'core'. The shell is a region whose connections are largely with the limbic system and which can be viewed as part of an 'extended amygdala', while the core has connections largely with the motor systems of the basal ganglia⁷. Put another way, these relationships mean that the shell may be a key link between midbrain and forebrain, playing a role in motivational and emotional functions, while the core is involved in the coordina-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Getting fixed up — a couple of smokers set to work on releasing dopamine from the shell region of the nucleus accumbens.

the conclusion that nicotine is indeed addictive. The authors show that in rats receiving small intravenous doses of nicotine — corresponding to doses which the animals will self-administer — there is a highly selective increase in the release of the neurotransmitter dopamine, and an activation of energy metabolism in the precise region of the forebrain that is similarly activated by cocaine, amphetamine and morphine.

The part of the brain involved is the nucleus accumbens, a small dopamine-rich brain region situated ventrally at the head of the striatum, which plays a crucial role in the rewarding effects of these addictive drugs⁴. Animals can be trained to self-administer such drugs directly into the nucleus accumbens, or into the ventral tegmental area in midbrain from which the dopaminergic fibres innervating the nucleus accumbens originate. Measurements of dopamine release made by means of microdialysis probes in the

of movements.

Last year, the Cagliari group reported that cocaine, amphetamine and morphine provoked a selective release of dopamine from the shell region of nucleus accumbens in rat brain⁸, and that this was accompanied by a similarly selective increase in energy metabolism in this brain region⁹. The present paper³ shows that nicotine mimics precisely these effects, with highly selective effects on the shell region of nucleus accumbens, thus placing it in the same category as the other psychostimulants and opiates in terms of its effects on a key structure in the brain circuits involved in drug reward and drug-seeking behaviour.

Many questions remain unanswered. What, for instance, are the consequences of activation of dopamine receptors by the drug-induced release of dopamine in the nucleus accumbens shell? Findings published earlier this year from a group at the University of Yale suggest that stimulation

of both D1 and D2 dopamine-receptor subtypes may be important¹⁰. Their findings in animals trained to self-administer cocaine implied that drugs that act to stimulate D1 receptors could partially substitute for cocaine, whereas drugs which stimulate D2 receptors appeared to provoke further drug-seeking behaviour. These results suggest that D1 dopamine agonists might have a therapeutic use in treating withdrawal from addictive drugs, while D2 antagonists might help to reduce craving.

The brain receptors targeted by nicotine are more clearly understood. More than 90 per cent of high-affinity binding sites for nicotine in mammalian brain are accounted for by a subtype of the acetylcholine nicotinic receptor containing α_4 - and β_2 -subunits¹¹. This neuronal nicotinic receptor differs from that found in muscle in displaying a much higher affinity for nicotine and related alkaloids.

The precise anatomical location of the nicotinic receptors that are responsible for its effects on nucleus accumbens, however, is not known. Both the nucleus accumbens itself and the ventral tegmental area possess a high density of nicotinic receptors, and nicotine may act either to stimulate activity in the ventral tegmental dopamine neurons or locally on dopamine nerve terminals in the nucleus accumbens to provoke dopamine release. Whether the accompanying increase in energy metabolism is triggered directly by the released dopamine is also not known, although this would seem likely, as the same change is seen in response to other psychostimulants that act to increase dopamine release.

In economically developed countries, about three million people die each year from smoking, half of them before the age of 70, and these represent about 20 per cent of all deaths. Cancers of eight different sites are directly associated with smoking, and six other potentially fatal diseases are also judged to be caused by smoking. There is no doubt that smoking kills, and attitudes to the habit have changed radically in the past 30 years.

For centuries following its introduction, however, tobacco smoking was regarded as neither harmful nor addictive. New scientific evidence in the 1950s and 1960s persuaded most people that smoking was harmful — but it was still not considered to be addictive, despite the high relapse rate in smokers who attempt to stop (more than 95 per cent within 12 months, similar to the rate for heroin addicts). By the 1980s, however, the evidence that nicotine is addictive became more and more persuasive, leading to the Surgeon General's Report in 1988 and the hard-hitting attack on nicotine addiction in young people by the head of the Food and

Drug Administration last year¹². That message is now reinforced with the latest evidence from the Cagliari group.

It seems that King James I of England was well ahead of his time, when, in 1604, he wrote of smoking¹³:

A custom loathsome to the eye, hateful to the nose, harmful to the brain, dangerous to the lungs and in the black, stinking fume thereof, nearest resembling the horrible Stygian smoke of the pit that is bottomless.

Leslie L. Iversen is in the Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK.

MATERIALS SCIENCE

The perfect nanotube

Philip Ball

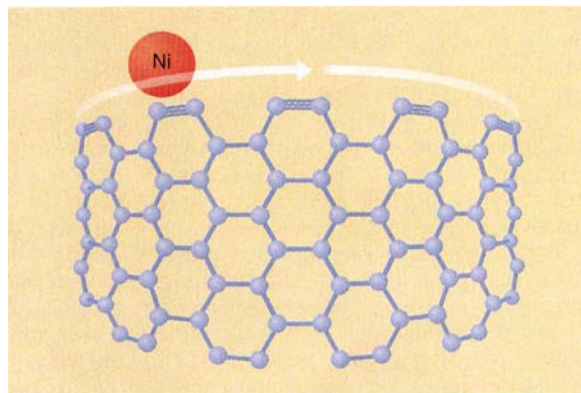
WHEN the first experimental signature of C_{60} appeared in the literature, it looked to be an unremarkable creature. In the mass spectrum of carbon clusters produced by laser ablation of graphite in 1984¹, it was merely one of a crowd — a peak amongst many other even-numbered clusters, a little larger than the others perhaps, but hardly evidence for the unique molecule that has since proved to be the Ur-fullerene.

By the same token, no one had thought to ask whether, amid the hordes of elongated fullerenes known as carbon nanotubes that have been described since their discovery² in 1991, there might be one that is special — a sort of Ur-nanotube whose status is to the others as that of C_{60} is to other near-spherical fullerene cages. But at a conference last week*, Richard Smalley of Rice University claimed to have found this very beast.

At first sight, it appears most unlikely that any particular nanotube should lay claim to a special status. Not only do conventional methods provide tubes of seemingly arbitrary diameter and any number of cylindrical graphitic shells nestled inside each other like Russian dolls, but for each tube there is an infinite choice of pitch of the ribbons of six-membered rings that wind around the walls in a helical fashion. Yet the experimental evidence for a privileged tube structure seems compelling. Smalley described fibrous carbon 'ropes' in which up to a hundred or so single-walled monodisperse

nanotubes (that is, having the same diameter) are packed together in a triangular lattice perfect enough to diffract, and formed from the carbon feedstock in yields of up to 90 per cent. Each of the tubes appears to have a diameter of 13.8 Å, and all extend from one end of the rope to the other³.

The ropes are deposited from the gas phase in a high-temperature oven in



The 'scooter' theory for how the (10,10) nanotube grows. The open mouth is lined with carbon-carbon triple bonds, and a nickel atom coordinated to the hexagon edges circulates around the rim of the tube, ensuring that defects such as pentagons (which would lead to tube closure) are annealed away.

which a graphite rod is vaporized by laser beams. The method is an adaptation of both the laser-ablation technique⁴ used for fullerene synthesis in 1985 and the metal-catalysed route⁵ to single-walled nanotubes introduced by researchers at IBM in 1993, as the graphite rod contains dispersed particles of a catalytic cobalt-nickel alloy. But a critical modification to the laser-ablation approach is to use a dual beam of coaxial lasers, one red and one green, so that the ablated carbon plasma gets a second dose of photons. At

1. United States Dept Health and Human Services *The Health Consequences of Smoking: Nicotine Addiction*, Rep. Surgeon Gen. Office on Smoking and Health (Government Printing Office, Washington DC, 1988).
2. Stolerman, I. P. & Jarvis, M. J. *Psychopharmacology* **117**, 2–10 (1995).
3. Pontieri, F. E., Tanda, G., Orzi, F. & Di Chiara, G. *Nature* **382**, 255–257 (1996).
4. Koob, G. F. *Trends pharmac. Sci.* **13**, 177–184 (1992).
5. Clarke, P. B. S. *Ciba Fdn Symp.* **152**, 152–163 (1990).
6. Imperato, A., Mulas, A. & Di Chiara, G. *Eur. J. Pharmac.* **132**, 337–338 (1986).
7. Heimer, L., Zahm, D. S., Churchill, L., Kalivas, P. W. & Wohltmann, C. *Neuroscience* **41**, 89–125 (1991).
8. Pontieri, F. E., Tanda, G. & Di Chiara, G. *Proc. natn. Acad. Sci. U.S.A.* **92**, 12304–12308 (1995).
9. Pontieri, F. E. et al. *Neuroreport* **5**, 2561–2564 (1994).
10. Self, D. W., Barnhart, W. J., Lehman, D. A. & Nestler, E. J. *Science* **271**, 1586–1590 (1996).
11. Whiting, P. & Lindstrom, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 595–599 (1987).
12. Kessler, D. A. *New Engl. J. Med.* **333**, 186–189 (1995).
13. James I *A Counterblast to Tobacco* (1604).

* Fullerenes '96, Oxford, UK, 7–10 July 1996.

the cool end of the reaction chamber, one finds a tangled 'felt' in which each strand is a bundle of single-walled nanotubes.

The crucial question is what feature of the nanotubes' structure engenders their monodispersity. Smalley faced an equivalent puzzle eleven years ago when contemplating a mass spectrum of carbon clusters that showed the 60-atom cluster to be (under the right conditions) overwhelmingly favoured over its neighbours⁴. And as before, it seems to be all a matter of where one places the six- and five-membered rings. Careful consideration of the problem has prompted the confident suggestion that the nanotubes all have an identical helicity: that which is denoted (10,10) (in a notation that characterizes the diameter and pitch in terms of the arrangement of six-ring ribbons). What is special about this configuration?

Most nanotube helicities give steps around the circumference of the open tube end, where, if you like, each helical ribbon terminates. But for the (10,10) tube, the open end has a crenellated rim in which each six-ring has an edge perpendicular to the tube axis. That is a good thing, Smalley suggests, because this arrangement uniquely allows the outermost edges of the 'exposed' six-rings to be capped with carbon-carbon triple bonds, leaving no dangling bonds at the tube end.

True, it means that each of these acetylenic groups will be uncomfortably strained, bent to about 120° rather than the customary 180°. But calculations suggest that the energetic price is not as high as one might imagine; and after all, the benzyne units that are the parent structures of these putative end rings are not unknown to organic chemists, even though they are more than a little reactive.

If the tube is to keep growing rather than closing off into the kind of short nanotubes seen in most previous syntheses, pentagonal rings (which curve the sheets inwards) have to be prevented. This (the story goes) is where the metal catalyst comes in. We are invited to imagine a single nickel atom scooting around the circumference of the open tube, annealing away defects such as five-membered rings as soon as they are formed. And why not?

Meanwhile, there is the need to account for the fact that all the tubes in a rope apparently run from one end to the other, while some must surely become attached to the bundle (in the gas phase) at a later stage than others. That can be explained if one assumes that short tubes deposited on long ones are fed by diffusion of the carbon feedstock (in whatever form that may take) along the surfaces of the long tubes, as well as by deposition from the gas phase, until the open end catches up with those of the other tubes.

Why, though, the magic number of

Ganymede: wrinkled but magnetic



ON 27 June, the Galileo spacecraft flew within 7,500 km of the surface of Ganymede, Jupiter's favourite boy (and largest moon). The first pictures taken were released last week, including this view of cratered and mountain-ridged crust, 55 km top to bottom. The high resolution of these images reveals more complicated tectonic processes than had been widely expected, including shearing like that on the San Andreas fault.

Another surprise was that Ganymede has a magnetic field — the first to be firmly detected around any moon. It may be generated by convection in a layer of salty water below the surface; if so, the same layer might also be involved in the surface tectonism.

We will learn a lot more in November, when Galileo meets Callisto, nearly Ganymede's twin in size and composition, if not in looks. Comparing the two worlds should help us to understand each one.

S. B.

13.8 Å, given that one can postulate larger tubes that also have the favourable crenellated ends if they have the right helicity? It turns out that 13.8 Å is (near enough) the diameter of a tube built up from a hemispherical cap of half a C₂₄₀ fullerene. The significance of that is that C₂₄₀ is the first fullerene larger than C₆₀ that shares its symmetry, with the 12 pentagons in the same icosahedral arrangement. For an ideal carbon onion built around a central C₆₀, C₂₄₀ is the next layer. The uncanny thing is that a cap of this sort makes the (10,10) nanotube just the right size to drop a C₆₀ molecule down the middle, because when C₆₀ sits symmetrically within C₂₄₀, the distance between the carbon shells is almost an exact match for the inter-layer spacing in graphite. This is a freak of nature. There is no earthly reason why the shell separation between the two polyhedra, which is set (in relation to, say, the hexagon width) by geometry alone, should so nearly coincide with the layer spacing of graphite, set by the strength of the van der Waals interaction. But there it is; the archetypal nanotube turns out to be a perfect pea-shooter for C₆₀.

The (10,10) nanotubes are among the one-third of all possible types whose atomic structure is predicted to predispose them to metallic conductivity. This expectation seems to be borne out: measurements of the conductivity of individual ropes, unravelled from the mass by the bargain-basement method of applying Scotch tape and pulling, suggest along-axis resistivities of less than 1.0 μΩ m (only twice the in-plane resistivity of graphite) (J. E. Fischer, Univ. Pennsylvania). This is probably an upper limit, but measurements on individual tubes will be trickier.

One hesitates to say it, but this could just be the breakthrough that nanotube science has been waiting for. When they were first reported⁵, single-walled nanotubes were hailed as a panacea for the problems of comparing observation and prediction, because their properties — mechanical, electronic and magnetic — would not be the convoluted product of those of several walls. But attempts to study them have been held back by the difficulty of obtaining samples free from other carbon and non-carbon debris (not to mention the range of tube diameters). Now nanotube preparation stands a chance of becoming less a matter of crude materials processing, with products sharing little more than composition, general morphology and inter-layer spacing, and more a case of macromolecular synthesis, yielding identical molecules that happen to be micrometres long. Indeed, the new method bears a startling resemblance to the 'living' polymerization techniques that have revolutionized polymer science: a metal atom chaperones the reactive end of a growing polymerized structure so as to ensure the controlled formation of defect-free chains of uniform length. And just as living polymerization has opened new doors in polymer technology, so these monodisperse nanotubes may become the strands of a nanotechnological materials science. □

Philip Ball is an associate editor at Nature.

1. Rohlfing, E. A., Cox, D. M. & Kaldor, A. *J. chem. Phys.* **81**, 3322 (1984).
2. Iijima, S. *Nature* **354**, 56–58 (1991).
3. Smalley, R. E. *et al. Science* (in the press).
4. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–163 (1985).
5. Bethune, D. S. *et al. Nature* **363**, 605–607 (1993).

The acid test for band 3

Michael J. A. Tanner

THE plasma membrane of the red blood cell has some 15 major protein constituents, one of them being band 3; this is a transmembrane protein, and gets its name from its relative position on gels after electrophoresis. Band 3 is central to the function of red cells, having two roles. It maintains the integrity and stability of the erythrocyte by anchoring the erythrocyte skeleton to the cell's lipid bilayer, and it increases the capacity of blood to carry CO_2 , as HCO_3^- , from the tissues to the lungs — it does this because it is an anion exchanger, and indeed is also known as AE1 (for anion exchanger 1). People who are partially deficient in band 3 are not uncommon in the population, and they suffer from a form of haemolytic anaemia, known as hereditary spherocytosis, in which the shape of the red cells is distorted. But the complete absence of erythrocyte band 3 has never been described before and the protein has been regarded as being essential for life.

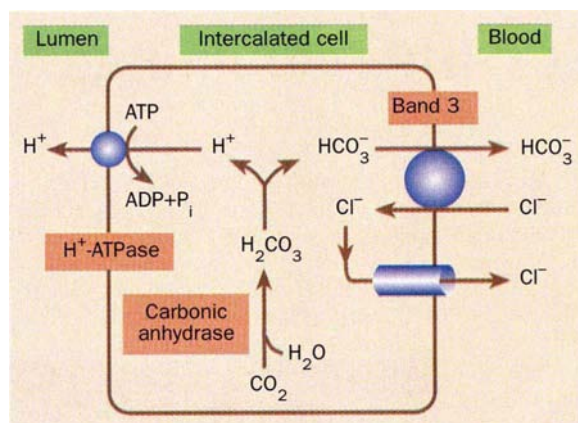
Two reports now challenge this view — Inaba and co-workers¹ have found Japanese black cattle that are completely deficient in erythrocyte band 3, and a preliminary report by Peters *et al.*² describes the production, by gene disruption methods, of mice that are likewise deficient. Some of these animals survive, although they have severe haemolytic anaemia and retarded growth.

Band 3 is expressed only in red cells and acid-secreting cells in the kidney (of which more later). The protein³ contains two domains: the amino-terminal portion binds the erythrocyte skeleton and the carboxy-terminal membrane domain carries out anion transport. The erythrocyte skeleton, a two-dimensional network of proteins located under the lipid bilayer, enables the cell to remain intact despite the physical stresses of the circulation⁴.

Hereditary spherocytosis⁴ manifests itself as osmotically fragile, small, red cells with varying degrees of spheroidicity. The disease results from defects in any of the several skeletal components that stabilize the lipid bilayer. Unsupported areas of bilayer bud off as microvesicles, and the resulting decrease in surface area/volume forces the shape of the erythrocyte from the usual biconcave disc towards a sphere. The spheroidal cells are less able to deform and squeeze through the narrow

passages in the spleen, so they become sequestered there.

Up to 25 per cent of cases of hereditary spherocytosis have a reduction of erythrocyte membrane band 3 (by 20–40%), because of the heterozygous presence of an altered band-3 gene. Several band-3 mutations are known (summarized in refs 4 and 5), and the mutant proteins are not expressed in the erythrocytes. The band-3-deficient cattle described by Inaba *et al.*¹



Function of band 3 in acid secretion by intercalated cells in the renal distal tubule¹⁰. The final portion of the kidney tubule — the distal tubule and collecting duct — secretes acid into the tubular fluid and also determines the final concentration of urine. This tubular epithelium has two main types of cell, the intercalated cells which secrete acid and the principal cells which regulate water permeability. The figure illustrates the mechanism of acid secretion by the intercalated cell. The apical H^+ -pumping ATPase in the cell drives protons into the tubular lumen, where they are buffered by phosphate and ammonia in the luminal fluid. The depletion of intracellular protons (which would otherwise alkalize the cell interior) is countered by a supply of protons derived from the intracellular hydration of CO_2 by carbonic anhydrase. This forms intracellular HCO_3^- which is transported into the blood in exchange for chloride by band 3, and the chloride ions that have moved into the cell leave through chloride-selective channels.

are homozygous for a nonsense mutation, and no traces of the predicted truncated band 3 are found in the red cells.

How does the complete absence of band 3 affect red cells? Less than one might anticipate. The adult animals had anaemia, with haematocrit (packed cell volume of blood) substantially reduced. The erythrocytes were almost spherical and markedly unstable, shedding small vesicles from their surfaces. The membranes retained the skeletal proteins, but

in reduced amounts (although band 4.2, which is known to attach to band 3, was absent)^{1,2}. Electron microscopy by both groups showed the presence of an assembled skeletal network, although it was somewhat disorganized. The number of intramembrane particles revealed by freeze-fracture was reduced by 50–70%, leaving large protein-free areas. Newborn animals had very low haematocrits because of haemolysis (3% and 13% for the mice and cattle, respectively), and the death rate was highest at this stage (88% of the mice, 45% of the cattle).

As expected, the band-3-deficient red cells showed very little anion transport.

CO_2 is converted to HCO_3^- within erythrocytes by carbonic anhydrase. HCO_3^- is normally equilibrated between the cells and the blood plasma by the $\text{Cl}^-/\text{HCO}_3^-$ exchange activity of band 3. However, in the band-3-deficient animals, HCO_3^- is restricted to the cells and this, together with the low haematocrit, allows only a small proportion of the blood volume to carry HCO_3^- . The resting adult band-3-deficient cattle (25% haematocrit) had mild acidosis — over-acidity — with decreased blood HCO_3^- , increased blood CO_2 and blood pH reduced by 0.15 pH unit. The last may account for the retarded growth of the animals observed by Inaba *et al.* The acidosis would increase with exercise: it must be especially severe in the newborn animals, because of their very low haematocrits, and may contribute to their high death rates.

Band 3 is also involved in regulating whole-body acid-base balance by the kidney. Net non-volatile acid clearance occurs by acidification of the urine, while net base clearance results in alkalization of the urine by secretion of HCO_3^- . Acid secretion by the α -intercalated cells of the distal nephron requires $\text{Cl}^-/\text{HCO}_3^-$ exchange (see figure), and this process is mediated by band 3 (ref. 6). Normally, calves can acidify their urine to pH 5.3–5.6 in response to an acid load⁷, but despite their acidosis, the band-3-deficient cattle could not acidify their urine to less than pH 7.5.

This defective renal-acid secretion also impairs the animals' ability to correct acidosis. In normal animals, diet largely determines whether there is a net non-volatile acid or base excess⁸. Carnivores and omnivores with a high protein diet excrete an acidic urine; herbivores excrete

1. Inaba, M. *et al.* *J. clin. Invest.* **97**, 1804–1817 (1996).
2. Peters, L. L. *et al.* *Blood* **86** (Suppl. 1), 124a (1995).
3. Tanner, M. J. A. *Semin. Hemat.* **30**, 34–57 (1993).
4. Lux, S. E. & Palek, J. in *Blood: Principles and Practice of Hematology* (eds Handin, R. L., Lux, S. E. & Stossel, T. P.) 1701–1817 (Lippincott, Philadelphia, 1995).
5. Delaunay, J. *FEBS Lett.* **369**, 34–37 (1995).
6. Alper, S. L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 5429–5433 (1989).
7. Scott, D. *et al.* *Q. J. exp. Physiol.* **56**, 18–32 (1971).

8. Houpt, T. R. in *Dukes' Physiology of Domestic Animals* (ed. Swenson, M. J.) 486–506 (Cornell Univ. Press, London, 1984).
9. Feest, T. G. & Wrong, O. M. in *Recent Advances in Renal Medicine* No. 2 (eds Jones, N. F. & Peters, D. K.) 243–271 (Churchill-Edinburgh, 1982).
10. Alpern, R. J. & Rector, F. C. in *Handbook of Physiology: Sect. 8, Renal Physiology* (ed. Windhager, E. E.) 767–812 (Oxford Univ. Press, 1992).

an alkaline urine, because the organic anions in their leafy diet produce base as HCO_3^- . The normal controls for Inaba and colleagues' band-3-deficient cattle had urine of pH 8.0 and were in base excess, and this dietary base excess may ameliorate the acidosis in the cattle lacking band 3. In humans, renal tubular acidosis is treated by sodium bicarbonate. The same approach might prove beneficial to the band-3-deficient animals, and might also improve the life expectancy of the newborn animals, which probably have an acid excess from their milk diet as well as respiratory acidosis.

Finally, the band-3-deficient animals should be valuable for studying mammalian acid-base physiology. They might also provide an animal model for human distal renal tubular acidosis — defective acid secretion in the distal tubule. This disease can originate in several ways and has a variety of clinical features, the most serious of which are kidney stones, bone disease and muscle weakness⁹. □

Michael J. A. Tanner is in the Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, UK.

GEOCHEMISTRY

Earthquakes with gold linings

Bruce Yardley

GOLD has been prized since the earliest times, not just because of its appearance, but because of its durability: it will not rust, and appears to be insoluble in our everyday experience. But if we look at the origins of the ore deposits from which gold is extracted, they pose something of a conundrum. In many cases, the gold occurs as grains of metal which have clearly precipitated from solution. Gold can be relatively mobile in soils, under certain conditions, but in the reducing conditions that prevail throughout most of the Earth it is effectively transported in solution as bisulphide complexes¹. There are a variety of geological settings in which appropriate sulphide-bearing ore fluids might be generated, and the origins of such fluids, and the factors that can trigger gold precipitation from them, have been subjects of lively controversy. A recent contribution² by Wilkinson and Johnston* has not only characterized the broad chemical composition of the gold ore fluids at the Croagh Patrick prospect, western Ireland, but has also demonstrated a link between the precipitation of gold in fractures and contemporary faulting.

Croagh Patrick is an example of a mesothermal gold deposit: the gold occurs in quartz veins cutting across weakly metamorphosed but strongly deformed sediments. These veins formed after the pervasive deformation of the rocks, while they were being uplifted and faulted, at temperatures of about 300 °C and depths of around five kilometres. Microscope freezing studies of fluid inclusions in vein quartz, trapped at the time that gold was also precipitating, typically show low salinities, and the inclusions usually contain a central bubble of liquid carbon dioxide that has unmixed from the water as it cooled to surface conditions.

The essence of forming a good hydrothermal gold deposit, from a geochemical point of view, is that a fluid must precipitate as much gold as possible, with as little as possible of any other minerals, in the smallest possible rock volume. Assuming

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASON
REASONS

Natural gold crystals on quartz. The two form together in veins that grow from hot aqueous solutions in the upper crust, and may result from pressure changes triggered by earthquakes.

the presence of hydrogen sulphide at modest levels (and these have occasionally been measured³), low-salinity $\text{H}_2\text{O}-\text{CO}_2$ fluids have the optimum potential to transport gold and quartz, but will be low in many other metals that are better complexed by chloride, or cannot remain in solution in the presence of much sulphide.

An important pointer as to why gold precipitates from these fluids comes from another major class of gold ore bodies, the epithermal deposits. These form in geothermal circulation systems around active

volcanoes^{4,5}. Gold precipitation occurs when the geothermal waters boil, because hydrogen sulphide is partitioned into the steam phase, thereby destabilizing gold bisulphide complexes present in solution in the liquid phase. So, many authors have suggested that in mesothermal deposits the gold is precipitated from the $\text{H}_2\text{O}-\text{CO}_2$ fluid as a result of unmixing owing to the onset of fluid immiscibility⁶⁻⁸ — a single-phase fluid of intermediate composition crosses the solvus and a vapour-like CO_2 fluid separates out, into which H_2S is partitioned.

Fluid inclusions in minerals can preserve evidence of immiscibility by trapping examples of each of the phases that formerly coexisted in the rock as separate inclusions, one richer in CO_2 , the other in H_2O . When the sample is heated under the microscope, both inclusions will homogenize to a single phase at a temperature close to that of their original trapping in the quartz. Mesothermal gold deposits provide such evidence for fluid immiscibility, which can arise either because a mixed-volatile fluid that formed in the single-phase region of pressure-temperature space then moves into conditions where two phases are stable, or because a fluid changes composition by reaction with wall rocks.

What Wilkinson and Johnston have done² is to look at the fluid inclusions from the point of view of their density as well as their chemistry, and to consider the possible effects of pressure fluctuations in fractures on mixed $\text{H}_2\text{O}-\text{CO}_2$ fluids, and the implications for gold solubility. Large earthquakes are often accompanied by changed patterns of fluid flow at the surface, and a number of studies have shown that vein minerals formed at depth in fault planes may contain fluid inclusions of very variable density. This could be the result of large temperature variations. But many workers believe that it is more likely to be due to fluctuations in fluid pressure in the fault plane between low, near hydrostatic, values (corresponding to periods when the fault is open and well drained), and values near to lithostatic pressure (that may reflect conditions as the fault plane seals up by mineral precipitation).

At Croagh Patrick, Wilkinson and Johnston found evidence for just such

GeoScience Features

*David Johnston died during fieldwork on the Mayo coast last year.

1. Seward, T. M. *Geochim. cosmochim. Acta* **37**, 379–399 (1973).
2. Wilkinson, J. J. & Johnston, J. D. *Geology* **24**, 395–398 (1996).
3. Yardley, B. W. D., Banks, D. A., Bottrell, S. H. & Diamond, L. W. *Mineralog. Mag.* **57**, 407–422 (1993).
4. Krupp, R. E. & Seward, T. M. *Econ. Geol.* **82**, 1109–1129 (1987).
5. Hedenquist, J. W. & Lowenstern, J. B. *Nature* **370**, 519–527 (1994).
6. Goldfarb, R. J., Leach, D. L., Rose, S. C. & Landis, G. P. *Econ. Geol. Monogr.* **8**, 363–375 (1989).
7. Robert, F., Boullier, A. M. & Firdaus, K. *J. geophys. Res.* **100**, 12861–12879 (1995).
8. Bowers, T. S. *Geochim. cosmochim. Acta* **55**, 2417–2434 (1991).
9. Diamond, L. W. *Am. J. Sci.* **290**, 912–958 (1990).

large pressure fluctuations in gold-bearing quartz veins. Because the $\text{H}_2\text{O}-\text{CO}_2$ solvus moves to higher temperatures with decreased pressure in the upper half of the crust, they conclude that a pressure drop due to faulting triggered unmixing of an originally miscible fluid, with consequent partitioning of H_2S into the CO_2 phase, and precipitation of gold.

Many examples of gold mineralization are associated with faults and fractures, which are commonly inferred to have acted as conduits for large fluid fluxes. The model of Wilkinson and Johnston implies a much more active role for the faults: it is seismic activity on them which led directly to the gold being deposited. Whether their model is universally applicable remains uncertain, because many workers have looked for associations between gold deposition and fluid immiscibility, and some have found that gold deposition occurred from a single-phase fluid⁹, possibly in response to a temperature drop or changes in oxidation-reduction conditions.

A fundamental problem with these deposits is where the fluids themselves have actually come from. They have been variously ascribed to the release of volatiles at depth from continuing metamorphism (but these reactions require so much heat that they release fluids too slowly and progressively), to magmatic sources (but in some cases magmatic rocks of the right age are apparently missing), or to deep penetration of surface-derived waters (but where do they get their CO_2 from?).

Historically, gold from mesothermal veins has played an important part in the economic development of the Old World, but there are still major uncertainties in our understanding of why they form. Croagh Patrick will probably not provide the final answers, because the mountain is an important religious site, and local opinion precludes the possibility of mining in the foreseeable future. □

Bruce Yardley is in the Department of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK.

in patients with diabetes, it may be of particular relevance to the prevention of late diabetic complications in people who have had the disease for some time.

Although the use of insulin provides a life-saving treatment for insulin-dependent diabetes, patients often experience diabetic complications including eye and kidney disease, and atherosclerosis². Many of these complications are thought to derive from the failure of daily doses of insulin accurately to track the intake of food and consequent increases in blood glucose. In the normal pancreas, β -cells sense a rise in glucose and secrete increased amounts of insulin. Cells throughout the body respond to insulin by increasing glucose uptake and conversion to storage forms such as fat and glycogen, thus maintaining blood glucose levels near a fixed set-point. Because injections of insulin once or twice a day do not exactly mimic this normal feedback mechanism, it is inevitable that levels of glucose will rise and fall more dramatically in diabetic patients and that there will be periods of extra exposure to the damaging effects of glucose. Similar complications are also observed in non-insulin-dependent diabetes, in spite of a variety of measures taken to control glucose levels³.

One of the best-characterized effects of excess blood glucose is the formation of protein adducts, a process known as glycation⁴ (see figure). Increased levels of glycation on proteins with modest half-lives reflect recent short-term increases in glucose levels. On more stable proteins, the initial glycation products become irreversibly fixed in a variety of chemical forms which have been given the evocative acronym of AGEs, or advanced-glycation end-products. The accumulation of AGEs represents an integrated measure of exposure to glucose over time. ►

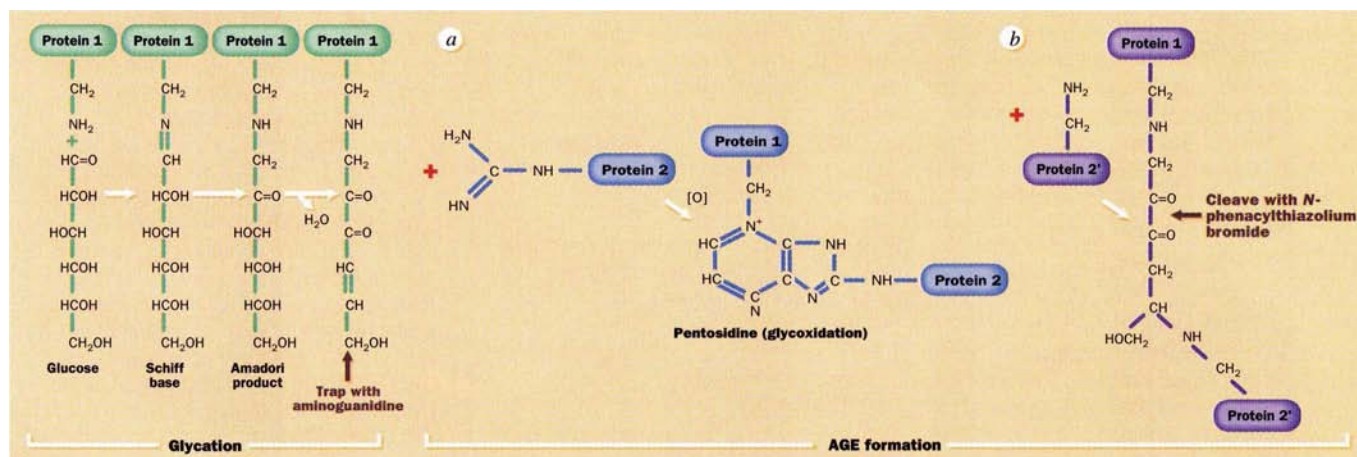
DIABETES

Breaking the curse of the AGEs

Kurt Drickamer

'RATIONAL drug design' is not a term that can be taken at face value. It implies that therapeutic agents emerge from a full molecular understanding of a disease or symptom — that drugs are designed to interact with biological targets for which structures and chemical mechanisms are known. Reality is not so clear cut. The development of treatments is much more likely to proceed hand in hand with incremental increases in basic biochemical knowledge.

A good example of this process is described on page 275 of this issue, where Vasan *et al.*¹ show how a partial understanding of a disease can lead to a potential approach for producing new drugs, and at the same time provide further understanding of the underlying biochemistry. Their results suggest a way to undo the crosslinking of proteins that occurs in all of us during a lifetime of exposure to blood sugar. As this process is accelerated



Chemical steps that may result in protein-protein crosslinking following glycation. One pathway (b) is postulated to follow directly from reaction of the diketone intermediate with amino groups of other proteins^{1,5}. For simplicity, ϵ -amino groups of lysine side chains are shown in the first and second steps, although α -amino groups might participate in the first step and a variety of protein side chains could be involved in the

final addition reaction. Alternative mechanisms of crosslink formation (a), involving condensation with arginine side chains and oxidation (glycoxidation) to form pentosidine, have been proposed^{6,7}. In both cases, much of the chemistry underlying the formation of crosslinks remains hypothetical. Points at which the formation of AGEs (advanced-glycation end-products) can be disrupted are indicated in grey.

AGE-modified proteins are believed to represent tissue damage at the molecular level. Establishing the chemical structures of these heterogeneous and sometimes labile compounds has proven difficult, but it is generally agreed that one effect of AGEs is protein crosslinking, which results when an initially glycated protein reacts with a second protein⁵⁻⁷. These products probably lie on an important route to diabetic complications. In one of the several proposed pathways, glucose-induced crosslinking of collagen increases tissue stiffness, which may lead directly or indirectly to kidney failure.

For all that it is incompletely understood, the glycation pathway nonetheless provides opportunities for tackling the damage process that occurs during normal and diabetic ageing. For example, aminoguanidine, an analogue of the side chain of arginine, can trap the initial glycation adducts, thus preventing AGE formation⁸. Aminoguanidine is an effective treatment for animal models of diabetic complication and is being evaluated in humans. But what of damage that has already occurred? This issue matters especially for non-insulin-dependent diabetic patients, who may have sustained glycation damage over a period of many years before their condition was diagnosed.

In the face of incomplete knowledge of the chemical structures of the crosslinked products of glycation, Vasan *et al.*¹ based their approach to tackling such damage on a reasonable estimation of what might be happening (see figure). The diagram depicts two plausible ways (*a* and *b*) in which an intermediate containing two adjacent carbonyl groups might complex with amino-acid side chains to form intermolecular crosslinks. In *b*, this diketone structure is still present in the final crosslink.

Working on this assumption, Vasan *et al.* went on to synthesize *N*-phenacylthiazolium bromide, a crosslink 'breaker' that can cleave the diketone bridge, which they first tested on small-molecule models. In further experiments, crosslinking of collagen in diabetic rat-tail tendon, a commonly used indicator of glucose-induced damage, could be largely eliminated by treatment *in vitro*. Finally, *in vivo* administration of the crosslink breaker to diabetic rats resulted in release of up to half of the immunoglobulin attached to the surface of circulating red blood cells; as AGE-immobilized immunoglobulins provide a convenient example of protein crosslinking that accumulates at elevated levels in diabetes, this last result provides evidence that the crosslink breaker is active under physiological conditions.

This set of experiments represents a major step forward in our chemical understanding of glucose-induced damage. To the extent that *N*-phenacylthiazolium bromide selectively cleaves diketones, one

clear implication of these findings is that much of the crosslinking observed in the aged and diabetic tissue must contain these structures. If the specificity of the reagent is substantiated in further experiments, then these results can be taken as evidence favouring one (*b*) of the two crosslinking mechanisms shown in the figure.

Because *N*-phenacylthiazolium bromide appears to reverse crosslinking in live rats, we now also have evidence that AGE damage is not irreversible. This reagent therefore provides at least the conceptual basis for the design of drugs to reverse damage that has already occurred. But there are important issues that must be considered when attempting to translate these findings into benefits for patients. Long-term toxicity of possible drugs is the most obvious, but more subtle problems, such as antigenicity of the cleavage products, would also have to be considered. Nonetheless, these studies provide real prospects for reversing the effects of abnormally high glucose levels in diabetic patients. These patients are at high risk for glucose-induced damage, so some treatment, even with side-effects, may be better than nothing at all.

Glucose-induced damage is not, of course, unique to diabetic patients: even at normal levels of blood glucose there will be some degree of glycation, and the resulting damage will accumulate over time. Does the prospect of reversing glucose-mediated damage imply that a 'cure' for ageing lies around the corner? Vasan *et al.* report preliminary findings suggesting that their crosslink breaker can disaggregate the amyloid deposits that are a hallmark of Alzheimer's disease. Unfortunately, however, the association of glycation-induced crosslinking with age-related disease processes has only relatively recently come under investigation⁹, and we lack a good diagnostic tool for defining patients at highest risk who could be targeted with appropriate medication. So there are high hurdles to be surmounted before reversal of glycation-induced crosslinking can be considered a potential strategy for ameliorating the effects of non-diabetic ageing. □

Kurt Drickamer is in the Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.

The highest spin

NATURE invented the principle of the helicopter. For millions of years, her rotating winged seeds, such as those of the sycamore tree, have spun sedately to the ground each autumn. If you could somehow feed spin into a sycamore seed, it would generate lift. It could fly and soar as the simplest possible powered aircraft.

Daedalus now plans to do it. Imagine, he says, a circularly polarized beam of microwaves. Its magnetic vector rotates clockwise (say) at the beam frequency. Now mix it with a beam polarized in the opposite sense, and with a frequency just 10 Hz lower. Its magnetic vector will rotate counter-clockwise at this lower frequency. The two vectors will sum to a resultant which rotates clockwise at the difference frequency of 10 Hz. A magnetic sycamore seed exposed to the mixed beam would follow this rotating field, and would spin at 10 Hz like a little induction motor. So it would fly.

The technology could be very simple. A steerable microwave system would launch an appropriately mixed beam upwards. Its frequency difference would be slowly increased; a bladed ferrite rotor in the beam would begin to spin and would soon take off. As it rose along the beam, it would reflect or re-radiate some of it back to its source. This return would be modulated by the blade frequency, giving a feedback signal to the control system.

Steering this elegant craft will be tricky. Daedalus hopes that by offsetting the beam and polarizing it elliptically in a specific direction, it will tilt the rotor in that direction via the resulting gyroscopic precession. If (like some sycamore seeds) the rotor is almost rotationally unstable, even a small imbalance might do the job. Once the control strategy has been perfected, however, the 'sycacopter' will be free to travel anywhere its beam can reach. It will be able to climb, descend or hover, and stay up indefinitely.

Cheap and simple, sycacopters will get everywhere. Hovering at high altitude, they will replace satellites and masts as relays or reflectors for broadcasting and mobile telephony. Carrying TV cameras (whose circular scan will be provided by their spin), they will let Big Brother snoop wherever he wishes. Sycacopters with sharpened blades and explosive charges will rise into the path of cruise missiles, while versions with polishing mops will clean high windows and the inaccessible statuary of ancient cathedrals. And tiny ones, buzzing with fury, will attack flies, mosquitoes and even swarms of locusts, under manual or radar control: a splendid new way of countering mankind's small and ancient enemies.

David Jones

1. Vasan, S. *et al.* *Nature* **382**, 275-278 (1996).
2. Nathan, D. M. *New Engl. J. Med.* **328**, 1677-1685 (1993).
3. Nathan, D. M. *Clin. invest. Med.* **18**, 332-339 (1995).
4. Bucala, R. & Cerami, A. *Adv. Pharmac.* **23**, 1-34 (1992).
5. Chen, H.-J. & Cerami, A. *J. Carbohydr. Chem.* **12**, 731-742 (1993).
6. Wolff, S. P., Jiang, Z. Y. & Hunt, J. V. *Free Radical Biol. Med.* **10**, 339-352 (1991).
7. Fu, M.-X., Knecht, K. J., Thorpe, S. R. & Baynes, J. W. *Diabetes* **41** (Suppl. 2), 42-48 (1992).
8. Brownlee, M. A. *Rev. Med.* **46**, 836-841 (1994).
9. Harrington, C. R. & Colaco, C. A. L. S. *Nature* **370**, 247-248 (1994).

The Exodus enigma

SIR — We have obtained six high-precision radiocarbon dates of cereal grains from Tell es-Sultan (Jericho), stratigraphically related to the end of the Middle Bronze Age (MB-IIC). The cereal samples belong to stratigraphic phase HAF.XII–XIII.iii.iii, the time of the final destruction and conflagration of Middle Bronze Jericho¹.

possible relationship between this eruption and the Exodus plague of “darkness that can be felt” (Exodus 10:21), was suggested by geologists, who found Minoan ash in sediments of the eastern Mediterranean Sea^{4–6} and Nile delta⁶.

Two averaged ¹⁴C data sets on short-lived materials from Akrotiri, measured at Oxford (3,357±21 yr BP)⁷ and Copenhagen

radical change in Canaan at the transition between the Middle and Late Bronze Ages, as expressed by Gonen: “The causes of destruction of the Middle Bronze sites have not been adequately clarified... In any case, hardly a single site escaped massive destruction”⁹.

We calibrated¹⁰ the radiocarbon dates of Jericho and Akrotiri, using a smoothed function (smoothing parameter $S = 1$)^{10,11} through a decadal dendrochronologically obtained data set (see figure). We included both the original (1986) and revised (1993) version¹² in our calculations^{10,11} because it is uncertain which one is preferable¹³.

The calibrated date for the destruction of MB-IIC Jericho is admittedly ambiguous, due to a mid-sixteenth century ‘wiggle’ in the calibration curves. However, the low-chronology option requires unsatisfactory stretching of the 45-year difference between the Santorini and Jericho ¹⁴C dates in relation to the 1628 BC dendro-event¹⁵. We prefer the (medium) high chronology, which is older than the XVIIIth Dynasty (see figure). A high-precision wiggle-matched 1,503-year dendrochronology from Turkey¹⁴ provides major new evidence for the tree-ring signal of 1628 BC as the date of the Minoan Santorini eruption¹⁵. If the latter eruption is associated with the Exodus plague of darkness^{4–6}, then the 40-year desert period (see figure) fits the high chronology of our calibrated ¹⁴C date for Jericho’s destruction, calibrated^{10,11} with the 1993¹² calibration curve, which suits the 1628 BC dendro-signal. If we consider only the Akrotiri radiocarbon results, the 40-year desert period fits the Jericho (medium) high-chronology ¹⁴C dates, calibrated with either curve.

Various theories and dates, ranging

Data from six cereal samples

Lab. no.	¹⁴ C date (yr BP)	δ ¹³ C (‰)	Sample no.	Material
GrN-18539	3,312 ± 14	–23.09	Jp.H.14.3; Hopf largest box	Grains, <i>Hordeum vulgare</i>
GrN-18542	3,288 ± 20	–23.54	Jp.H.10.1b; Hopf #89	Grains, <i>Triticum</i> sp.
GrN-18543	3,331 ± 18	–23.31	Jp.H.10.1; Hopf #87	Grains, <i>Triticum</i> sp.
GrN-18544	3,312 ± 15	–23.28	SA-143 Jp.H.10.1 Hopf #87	Grains, cereal, fragmented
GrN-19063	3,240 ± 18	–23.72	SA-58 Jp.H.14.3g Hopf #92	Grains, <i>Hordeum vulgare</i>
GrN-19064	3,375 ± 25	–21.86	SA-36 Jp.H.14.3b Hopf #91	Grains, cereal, fragmented

The samples come from squares H-II, III and -VI at Tell es-Sultan (Jericho), belonging to stage/phase HAF.XII–XIII.iii.iii, (the end of the Middle Bronze Age). Sample numbers from the excavations reported in refs 1,2 are also included.

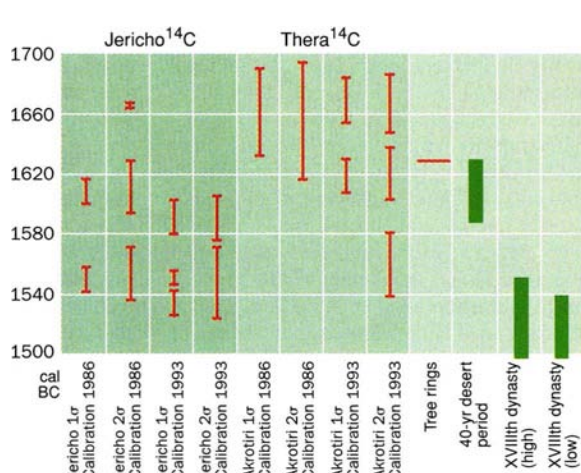
These cereal samples, excavated in the 1950s¹, were kindly provided by M. Hopf, who has investigated their palaeobotany². Conventional radiocarbon dating of these and other Jericho samples from the period was performed at Groningen University³. The table shows the six high-precision ¹⁴C dates. GrN-18539, -18542, -18543 and -18544 cluster outstandingly, averaging at 3,311 ± 8 years before present (BP). Inclusion of the oxalic acid standard σ in the last result gives a ¹⁴C date for Jericho’s destruction of 3,311±13 yr BP.

We note that the Minoan eruption at Santorini produced the largest volcanic dust cloud over the eastern Mediterranean in the second millennium BC. A

(3,355±32 yr BP)⁸, exist for the Minoan Santorini eruption. These averages, taken together, yield 3,356±18 yr BP, 45 radiocarbon years older than our ¹⁴C destruction date for MB-IIC Jericho. This time difference is rather striking, as it could fit the desert period of 40 years separating the Exodus from the destruction of Jericho, mentioned in ancient Hebrew texts.

Taking the geologically based Santorini–Exodus link as a serious option may be anathema for archaeologists who ascribe the widespread destruction of MB-IIC sites in Canaan, including Jericho, to Egyptian armies of the XVIIIth Dynasty. However, the latter association is unsatisfactory in social terms, considering the

Calibrated (smoothing parameter $S = 1$)^{10,11} radiocarbon dates of the destruction of MB-IIC Jericho, compared with calibrated (smoothing parameter $S = 1$)^{10,11} radiocarbon^{7,8} and tree-ring^{14,15} dates associated with the Minoan Santorini (Thera) eruption. The traditional 40-year period of the post-Exodus desert sojourn is tentatively indicated, as well as the high/low historical chronology for the beginning of the XVIIIth Dynasty in Egypt. Our calibration results for the destruction of MB-IIC Jericho are as follows. The 1986 data set¹²: (1 σ) 1617–1601 or 1559–1543 calibrated BC; (2 σ) 1668–1665, 1630–1595 or 1571–1538 cal BC; 1993 data set¹²: (1 σ) 1603–1580, 1556–1547 or 1543–1527 cal BC; (2 σ) 1606–1576 or 1572–1524 cal BC. For the Minoan Santorini eruption 1986 data set¹²: (1 σ) 1691–1633 cal BC; (2 σ) 1695–1617 cal BC; 1993 data set¹²: (1 σ) 1685–1654 or 1631–1607 cal BC; (2 σ) 1687–1648, 1638–1603 or 1581–1539 cal BC.



1. Kenyon, K. M. *Excavations at Jericho* Vol. 3 (Br. School Archaeol. Jerusalem, London, 1981).
2. Hopf, M. in *Excavations at Jericho* Vol. 5 (eds Kenyon, K. M. & Holland, T. A.) 576–621, Pl. 25–31 (Br. School Archaeol. Jerusalem, London, 1983).
3. Bruins, H. J. & Van der Plicht, J. *Radiocarbon* **37**, 213–220 (1995).
4. Ninkovich, D. & Heezen, B. C. *Colston Pap.* **17**, 413–464 (Butterworth, London, 1965).
5. Vitaliano, D. B. *Legends of the Earth: Their Geologic Origins* (Indiana Univ. Press, Bloomington, 1973).
6. Stanley, D. J. & Sheng, H. *Nature* **320**, 733–735 (1986).
7. Housley, R. A., et al. in *Thera and the Aegean World III* Vol. 3 (eds Hardy, D. A. & Renfrew, A. C.) 207–215 (Thera Fdn, London, 1990).
8. Friedrich, W. L., Wagner, P. & Tauber, H. in *Thera and the Aegean World III* Vol. 3 (eds Hardy, D. A. & Renfrew, A. C.) 188–196 (Thera Fdn, London, 1990).
9. Gonen, R. in *The Archaeology of Ancient Israel* (ed. Ben-Tor, A.) 211–257 (Yale Univ. Press, New Haven, 1992).
10. Van der Plicht, J. *Radiocarbon* **35**, 231–237 (1993).
11. Talma, A. S. & Vogel, J. C. *Radiocarbon* **35**, 317–322 (1993).
12. Stuiver, M. & Becker, B. *Radiocarbon* **28**, 863–910 (1986); **35**, 35–65 (1993).
13. Van der Plicht, J. & McCormac, F. G. *Radiocarbon* (in the press).
14. Kuniyoshi, P. et al. *Nature* **381**, 717–816 (1996).
15. Baillie, M. G. L. in *Thera and the Aegean World III* Vol. 3 (eds Hardy, D. A. & Renfrew, A. C.) 160–166 (Thera Fdn, London, 1990).
16. Cohen, R. *Biblical Archaeol. Rev.* July/August, 16–29 (1983).
17. Bruins, H. J. & Mook, W. G. *Radiocarbon* **31**, 1019–1029 (1989).

from around 2200 to 1200 BC, have been proposed for the Exodus tradition and emergence of Israel in Canaan, based on different interpretations of the archaeological dataset (ref. 16). Nevertheless, although they are powerful tools, archaeology and pottery are not the sole avenues that can be used to unravel the human past. Environmental events, high-precision radiocarbon dating and precise regional dendrochronologies may provide information unobtainable through archaeological associations¹⁷. New, high-precision ¹⁴C data series of XVIIIth Dynasty Egypt and Levantine Middle Bronze Age

sites are required. Wiggle-matching linkages with the Aegean dendrochronology¹⁴ may lead to truly absolute dating.

Hendrik J. Bruins

Jacob Blaustein Institute for Desert Research, Social Studies Center, and Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Sede Boker Campus, 84990 Israel

Johannes van der Plicht

University of Groningen, Centre for Isotope Research, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Marine gamete-release pheromones

SIR — In contrast to the large body of experimental evidence for the existence of sex pheromones in marine invertebrates, only a few chemical structures of these compounds have been elucidated^{1,2}. Here we report the isolation and identification of two gamete-release pheromones of marine invertebrates. By bioassay-guided fractionation and purification, we identified uric acid as the sperm-release pheromone of *Platynereis dumerilii* and inosine as the main component of the egg-release pheromone complex of *Nereis succinea* (Annelida, Polychaeta). These compounds contain the purine ring system, which has not as yet been identified in sex pheromones of any species.

Within the marine invertebrates, the Nereidae are well suited for chemical investigation of sex pheromones because of their distinctive breeding behaviour. In both species, reproduction strictly depends on an exact spatial and temporal synchronization. High population densities within restricted areas and the synchronization in time are responsible for the formation of swarms.

The reproduction process occurs as follows: (1) in these swarms mate recognition is achieved by chemical signals³. (2) At the moment of recognition the male discharges the egg-release pheromone (ERP)^{4,5}. (3) The female is stimulated to swim with high velocity in narrow circles surrounded by swarming males. After an

induction period of 10–40 seconds the female spawns. The discharged egg cloud contains the sperm-release pheromone (SRP). (4) Males achieve immediate fertilization by circling around the eggs and emitting sperm clouds.

It is known that the coelomic fluid of the opposite sex is the source of the gamete-release pheromones^{6,7}. With reference to the well-examined volatile sperm attractants in marine brown algae⁸, we first investigated volatile compounds of the coelomic fluid. These studies led to the elucidation of substances responsible for mate recognition in both species^{2,3,9,10}. However, none of these compounds elicited the gamete release, in contrast to the natural pheromone-containing coelomic fluid. Hence, in the present study we analysed the non-volatile, water-soluble compounds of the coelomic fluid for pheromonal activity.

To follow the target substances during isolation and purification, we developed a reliable behavioural bioassay. We injected samples with a microlitre syringe just in front of a mature specimen swarming in a glass dish. The release of a visible sperm cloud or loss of the total amount of eggs represented a positive pheromone response. We tested chromatographic fractions of the coelomic fluid both singly and in combination. (Details of methods are obtainable directly from the authors.)

In the case of *P. dumerilii*, we identified the bioactive single peak fraction finally obtained as uric acid. This SRP is effective at a threshold concentration of 0.6 μ M (solubility of uric acid in sea water, 32%, pH 8.2, 23 °C: 180 μ M). We measured the total amount of uric acid in the coelomic fluid of one female as 1.0 ± 0.1 μ g. When we dissolved this quantity of uric acid in 2 cm³ sea water (egg cloud volume) we found an actual pheromone concentration of 3.0 ± 0.3 μ M. This five-fold excess is essential to ensure successful fertilization taking into account natural conditions such as the effects of dilution by turbulence and currents in

sea water.

In the case of *N. succinea*, we identified inosine as the main component of the ERP-complex. (We are investigating the complete composition of the ERP complex separately.) We are now trying to identify the SRP of *N. succinea* and the ERP of *P. dumerilii*, so that we can compare males and females in the same species. Our preliminary results suggest that both uric acid and inosine are species- and sex-specific. They represent the first identified gamete-release pheromones in marine invertebrates. Remarkably, the purine ring system has not as yet been reported to be a feature in sex pheromones.

Erich Zeeck

Institute for Chemistry and Biology of the Marine Environment,

Tilman Harder

Manfred Beckmann

Carsten T. Müller

Department of Chemistry, Carl von Ossietzky University of Oldenburg, 26111 Oldenburg, Germany

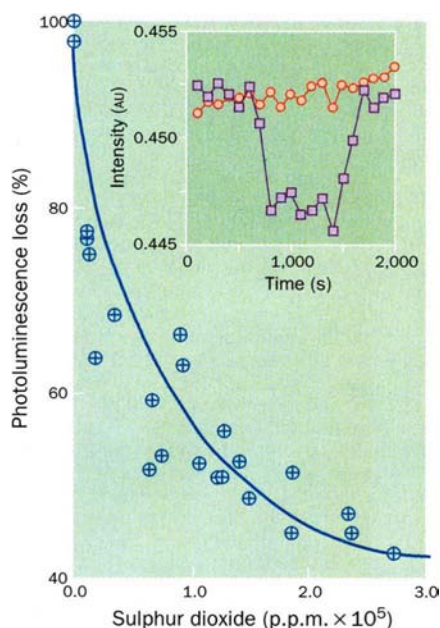
A silicon sensor for SO₂

SIR — Initially synthesized about 30 years ago, porous silicon is now attracting considerable attention because it both photoluminesces¹ and electroluminesces² in the visible portion of the optical spectrum. Although much of the recent interest in porous silicon is related to the possibility of generating an electroluminescent device having applications in the areas of display technology and information storage, we here suggest that the photoluminescent properties of this material make it attractive as a highly sensitive sensor for sulphur dioxide. Our suggestion is based on our recent observation that the photoluminescence of oxide-coated porous silicon is rapidly and reversibly quenched on a selective basis by gas-phase SO₂.

It is well documented that exposing freshly etched porous silicon to various organic reagents results in non-selective quenching of the observed photoluminescence³. We have also previously reported selective quenching by gas-phase and aqueous Brönsted bases when a thin surface oxide is present⁴. The observed quenching is reversible; subsequent exposure to a Brönsted acid restores the photoluminescence to its initial intensity. Partial quenching of the photosensitive luminescence by Brönsted bases results in a blueshifting of approximately 40 nm in the emission maximum. As with the photoluminescence intensity, exposure to acid returned the emission maximum to its initial position. The base-quenching/acid-restoration cycle can be repeated many times.

We now report that SO₂ gas is another effective quencher of oxidized porous sili-

- Kittredge, M., Terry, M. & Takahashi, F. T. *Fish. Bull.* **69**, 337–343 (1971).
- Zeeck, E., Hardege, J. D., Bartels-Hardege, H. & Wesselmann, G. *J. exp. Zool.* **246**, 285–292 (1988).
- Zeeck, E., Hardege, J., Bartels-Hardege, H. *Mar. Ecol. Prog. Ser.* **67**, 183–188 (1990).
- Boilly-Marer, Y. *Mar. Biol.* **24**, 167–169 (1974).
- Lillie, F. R. & Just, E. *Biol. Bull.* **24**, 147–160 (1913).
- Townsend, G. *Biol. Bull.* **77**, 306–307 (1939).
- Boilly-Marer, Y. in *Advances in Invertebrate Reproduction* Vol. 4 (eds Porchet, M., Andries, J. C. & Dhainaut, A.) 494 (Elsevier, Amsterdam, 1986).
- Jaenicke, L., Müller, D. G. & Moore, R. E. *J. Am. chem. Soc.* **96**, 3324–3326 (1974).
- Zeeck, E., Hardege, J. D., Bartels-Hardege, H. & Wesselmann, G. *Tetrahedron Lett.* **31**, 5613–5614 (1990).
- Zeeck, E., Hardege, J. D., Bartels-Hardege, H., Willig, A. & Wesselmann, G. *J. exp. Zool.* **260**, 93–98 (1991).



Photoluminescence intensity versus SO_2 concentration for porous silicon samples irradiated at 400 nm and monitored at 650 nm. Inset, An argon sulfur dioxide mixture (890 p.p.b. SO_2) flows through cell containing oxidized luminescent porous silicon. Emission intensity shows a significant drop on introduction of the SO_2 dose (time, 600 s) and subsequent restoration after the SO_2 flow is shut off (time, 1,400 s).

con. Sulphur dioxide is unique in several ways. It is a Lewis acid, and that alone makes it stand out from the other quenchers. Sulphur dioxide quenching does not require an acid to restore the loss of photoluminescence. To restore light emission to its original intensity, it is necessary only to remove the source of SO_2 . The blueshift of the photoluminescence maximum present in Brönsted base quenching does not appear, rather an obvious redshift occurs during SO_2 quenching⁵. Plots of photoluminescence intensity versus time during exposure to SO_2 and to a Brönsted base show kinetics that are more rapid for quenching by SO_2 ⁵. The differences observed between base quenching and SO_2 quenching indicate that the silicon surface- SO_2 interaction is distinct from the previously reported Brönsted base interaction.

We have used illumination from an argon-ion laser (458-nm line $<5 \text{ mW cm}^{-2}$ intensity) to irradiate porous silicon in a

quartz cell. On flowing a SO_2 /argon mixture through the cell, the photoluminescence intensity decreases. When we used $>10\%$ SO_2 /argon doses, we clearly saw luminescence quenching by the naked eye. Using a spectrofluorimeter to detect smaller changes in photoluminescence intensity, we found the limit of detection of SO_2 in argon is 440 parts per billion (p.p.b.). This concentration produces a reproducible 1.2% change in the observed photoluminescent intensity. Different instrumentation is expected further to reduce the limit of detection. We observed saturation above 70% SO_2 atmosphere, giving a useful working SO_2 concentration range from approximately 400 to 700,000 p.p.m. (see figure).

Exposure of the SO_2 -active oxidized porous silicon surface to high concentrations of the typical atmospheric and stack gases (O_2 , CO , CO_2 , H_2S and NO_x) does not illicit a quenching response. This makes it an interesting candidate for SO_2 monitoring with no pre-separation of samples necessary. At present, the surface reactivity toward SO_2 lasts for several weeks. Over that period of time, the layer of oxidized silicon grows thicker and the quenching response grows weaker, until eventually exposure to SO_2 results in no quenching. Experiments involving silyla-

tion of the surface show that porous silicon retains SO_2 reactivity on covalent bonding of a siloxy layer. Because such treatments are known to inhibit the oxidation of silicon surfaces⁶, it is probable that chemistry of this type can be used substantially to increase sensor lifetime.

In summary, we have observed selective gas-phase quenching of porous emission by SO_2 . The photoluminescence loss is reversible and the signal achieved is of analytical value. Preliminary experimentation with surface modification to lengthen the active lifetime of SO_2 sensitivity is promising. The high selectivity and sensitivity of porous silicon for SO_2 detection suggests possible applications in environmental monitoring. The mechanism for SO_2 quenching is apparently different from the better understood base-quenching mechanism, and requires further study. Clarification of the SO_2 -quenching mechanism is expected to lead to optimal sensitivity and lifetime, and may suggest new applications for porous silicon as a sensor.

Michael T. Kelly

Jonathan K. M. Chun

Andrew B. Bocarsly

Department of Chemistry,

Frick Laboratory,

Princeton University,

Princeton, New Jersey 08544-1009, USA

Species fragmentation or area loss?

SIR — Tilman *et al.*¹, extending a metapopulation model to include multi-species competition and habitat destruction, arrive at the paradoxical conclusion that, in a patchy habitat, the species most at risk to deterministic extinction when habitat is lost is the dominant competitor.

This result arises from the equilibrium between the dominant species' colonization of new sites occupied by inferior competitors and its mortality loss. The fraction of sites p occupied by the dominant competitor is described by the equilibrium relationship $dp/dt = 0 = cp(1-p-D) - mp$, where c and m are coloniza-

tion and mortality rates and D is the fraction of sites destroyed. The factor $(1-p-D)$ thus represents the proportion of sites available for colonization by the dominant competitor.

In an undisturbed habitat ($D = 0$), the initial equilibrium value of p is given by $p_0 = 1 - m/c$; in the disturbed habitat, $p = 1 - m/c - D = p_0 - D$. Thus if D is equal to or greater than p_0 , the dominant species is subject to deterministic extinction. A dominant species occupying 10% of an undisturbed habitat will be wiped out if as little as 10% of that habitat is destroyed. Habitat loss reduces effective colonization

by all competitors, but the dominant competitor most of all, as it relies on a trade-off between a low intrinsic colonization rate c and a large proportion of available sites to colonise, $(1-p)$ in the undisturbed condition.

What may be overlooked, however, is that the result obtains only if the sites selected for destruction are randomly distributed throughout the habitat, so that the same proportionate loss D of available colonization sites takes place in the local vicinity of every site occupied by the dominant

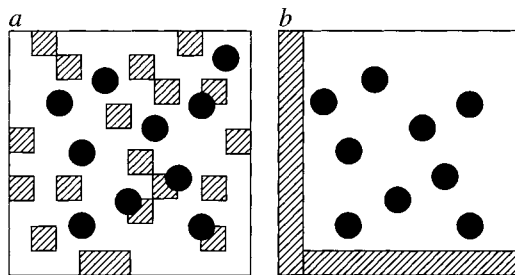


FIG. 1 a, Habitat destruction (shaded squares) randomly distributed throughout a patchy habitat proportionately reduces the sites open to colonization by the dominant competitors (black circles), reducing that species' effective colonization rate and, if the proportion of sites destroyed exceeds the proportion occupied by the dominant species, resulting in its eventual extinction. b, Destruction concentrated at the edge leaves the effective colonization rate unaffected.

1. Canham, L. T. *Appl. Phys. Lett.*, **57**, 1046-1048 (1990).
2. Richter, A., Steiner, P., Kozlowski, F. & Lang, W. *IEEE Electron Dev. Lett.*, **12**, 691 (1991).
3. Lauerhaas, J. M., Credo, G. M., Heinrich, J. L. & Sailor, M. J. *J. Am. Chem. Soc.*, **114**, 1911-1912 (1992).
4. Chun, J. K. M., Bocarsly, A. B., Cotrell, T. R., Benziger, J. B. & Yee, J. C. *J. Am. Chem. Soc.*, **115**, 3024-3025 (1993).
5. Bocarsly, A. B., Chun, J. K. M. & Kelly, M. T. in *Interfacial Design and Chemical Sensing* (eds Mallouk, T. E. & Harrison, D. J.) 78-90 (American Chemical Society, Washington DC, 1994).
6. Wrighton, M. S. *et al. J. Am. Chem. Soc.*, **100**, 1602 (1978).

species (a in the figure). The analytical model of Tilman *et al.* implicitly assumes this spatial homogeneity.

Should the destruction of habitat be concentrated in a contiguous block along the perimeter, however, then the undestroyed sites occupied by the dominant species will each remain surrounded in their immediate vicinity by the same proportion of available colonization sites as before the habitat loss occurred (b in the figure), and the effective colonization rate will remain unchanged. Indeed, the very fact that the equilibrium condition and extinction predictions of the analytical model are independent of the absolute size of the original habitat force this conclusion. Without fragmentation, the equilibrium condition of a 1-million-acre forest will be unchanged from that of a geometrically similar, "intact" 10-million-acre forest.

Thus if destruction is concentrated at the perimeter, the analytical model would predict no deterministic extinctions through competitive effects. In such a case, extinctions would only occur when the absolute area of the habitat is reduced to the point that the populations become small enough to fall victim to stochastic extinction processes².

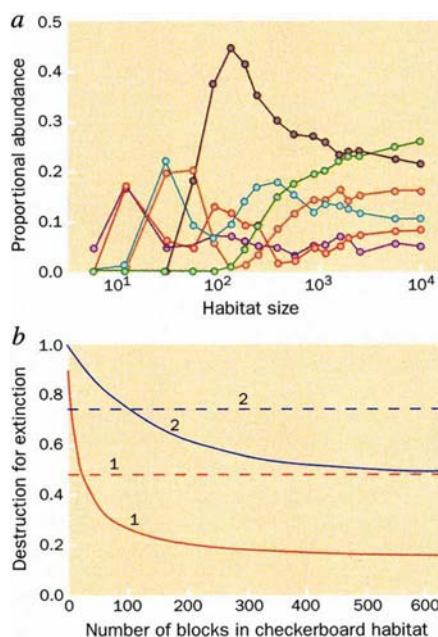
Extinction predictions based on species-area relations have been criticized^{3,4} for assuming a functional relation between area loss and species loss where in fact none may exist. Metapopulation models offer a convincing mechanism to predict species loss in highly fragmented habitats, but offer no support for the notion that area loss *per se* results in extinction. Predictions of global extinction rates based on estimates of area loss alone seem to be unreliable, unless it can be shown that permanent habitat destruction (and not just temporary forest clearing) is occurring in a highly random fashion that fragments the habitat consistently on a very local scale.

Stephen Budiansky

Black Sheep Farm, 14605 Chapel Lane,
Leesburg, Virginia 20176 USA

TILMAN *ET AL.* REPLY — Budiansky raises an important issue — that the quantitative relationship between the amount of habitat destroyed and the number of resulting species extinctions may depend on the spatial pattern of habitat destruction. Our analytical model¹ has no provision for habitat shape or size and cannot directly address this issue. However, a modified analytical model and extensive simulations of explicitly spatial versions of our model⁵ reveal two important effects.

First, the novel prediction of our analytical model¹ holds: for all spatial patterns of habitat destruction we tried, destruction leads to the biased extinction of superior competitors⁵. Even when they originally were the most abundant species,



superior competitors become extinct at lower levels of destruction than inferior competitors when an otherwise pristine habitat is reduced in size, such as by the loss of edge (Fig. 2). This biased extinction is a deterministic effect, caused by poor dispersal abilities, not the stochastic effect alluded to by Budiansky.

For habitats as large as those containing 104 sites, loss of edge differentially harms the best competitor. As edge is lost, the best competitor goes extinct first (species 1; Fig. 2a) in habitats of the size that Budiansky (his Fig. 1) suggested would have no edge effects. Further reductions in habitat size causes progressive extinction of species in order of competitive abilities (Fig. 2a), as occurred with destruction in our analytical model¹. Simulations show that better competitors, because of poorer dispersal, require larger areas for survival. The underlying reason for the robustness of our conclusions is the broad assumption that inferior competitors persist by virtue of greater dispersal ability and/or lower mortality rates.

It follows that better competitors, because of poorer dispersal, have larger minimal demands for habitat size. Furthermore, similar predictions of biased deterministic extinction of superior competitors arose when we modified our analytical model¹ by adding a term for the proportion of propagules lost to edges⁵ (via decreases in colonization rate proportional to $S^{-1/2}$ for square habitats, with S being habitat size). Other simulations showed that patch shape also influenced extinctions: for patches of a given area, those with greater perimeters caused more extinctions, with extinctions again biased toward the best competitors.

Second, additional simulations showed that extinctions can occur at amounts of habitat destruction markedly greater than,

markedly less than, or similar to those predicted analytically¹, depending on the spatial pattern of destruction (Fig. 2b). Extinctions occur at less destruction than analytically predicted if uniformly spaced blocks and/or small blocks of habitat are destroyed, and at more if large and/or clumped blocks are destroyed. Destruction of random small patches led to results like those predicted by our analytical model. In all cases, there is biased extinction of abundant, superior competitors.

These spatially explicit versions of our analytical model demonstrate that control of the spatial pattern of habitat destruction can be a significant management tool for minimizing extinctions. Equally as important, they show that habitat destruction can differentially harm abundant, superior competitors that are poor dispersers. This novel and alarming prediction of biased extinction was the central point of our paper. Its robustness, in response to many alternative spatial patterns of destruction, reaffirms our warning that the time-delayed extinction of originally abundant species may be an unexpected consequence of habitat destruction.

These spatially explicit versions of our analytical model demonstrate that control of the spatial pattern of habitat destruction can be a significant management tool for minimizing extinctions. Equally as important, they show that habitat destruction can differentially harm abundant, superior competitors that are poor dispersers. This novel and alarming prediction of biased extinction was the central point of our paper. Its robustness, in response to many alternative spatial patterns of destruction, reaffirms our warning that the time-delayed extinction of originally abundant species may be an unexpected consequence of habitat destruction.

David Tilman, Clarence Lehman

Department of Ecology, Evolution and Behavior,
University of Minnesota,
St Paul, Minnesota 55108-6097, USA

Robert May, Martin Nowak

Department of Zoology,
Oxford University, Oxford OX1 3PS, UK

1. Tilman, D. *et al.* *Nature* **371**, 65–66 (1994).
2. Soule, M. E. (ed.) *Viable Populations for Conservation* (Cambridge Univ. Press, 1987).
3. Budiansky, S. *Nature* **370**, 105 (1994).
4. Heywood, V. H. & Stuart, S. N. in *Tropical Deforestation and Species Extinction* (eds Whitmore, T. C. & Sayer, J. A.) 91–117 (Chapman & Hall, London, 1992).
5. Tilman, D. & Lehman, C. in *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions* (eds Tilman, D. & Kareiva, P.) (Princeton University Press, in the press).

Lives not worth living

Diane B. Paul

The Black Stork: Eugenics and the Death of "Defective" Babies in American Medicine and Motion Pictures since 1915. By Martin S. Pernick. Oxford University Press: 1996. Pp. 295. \$29.95, £22.95.

In 1915, Anna Bollinger of Chicago gave birth to a baby boy with multiple, severe physical abnormalities. Her physician, Harry Haiselden, persuaded the parents to allow their infant to die. The "Baby Bollinger" case would be followed by others in which Haiselden admitted to withholding treatment from newborns whose lives might have been saved by surgery. In his opinion, their remaining physical and/or mental deformities would result in lives that were not worth living.

Allowing severely impaired newborns to die was hardly a new practice, but these cases were exceptional in at least two respects: the active role of the doctor in persuading the parents, and the strenuous efforts he made to publicize rather than to conceal his actions. Haiselden displayed the dying infants to reporters and allowed the press access to the still-hospitalized mothers. He wrote newspaper articles and lectured to the public. He also helped to write and actually starred in a feature film, *The Black Stork*, dramatizing the Bollinger case. In every way he could, Haiselden strove to convince the public that withholding treatment from certain newborns was in the best interests of society and of the infants themselves.

In his fascinating account of this episode and its aftermath, Martin Pernick points out that controversies about euthanasia and refusal to treat impaired newborns have since erupted periodically. Each time, the issues are described as unprecedented products of new medical technologies. Hardly anyone now remembers Haiselden's once-famous campaign. We have come to associate eugenic euthanasia exclusively with Nazi Germany. That erasure of memory is unfortunate. The Bollinger case provides a window on little-discussed attitudes toward infants, medicine and the disabled in Progressive-era America. It raises thought-provoking questions about the history of eugenics and current controversies about the morality of allowing impaired newborns to die (such as in the 1982 Baby Doe case).

The number and ideological diversity of Haiselden's supporters are perhaps the book's most striking findings. The doctor was defended by nativists and immigrants, feminists and misogynists, socialists and businessmen. Roman Catholics of course were appalled, but many medical leaders

were also critical. They did not welcome this extension of their power. Most of those who favoured letting impaired newborns die felt that the decision should be left to physicians. Helen Keller, who was herself both deaf and blind and who

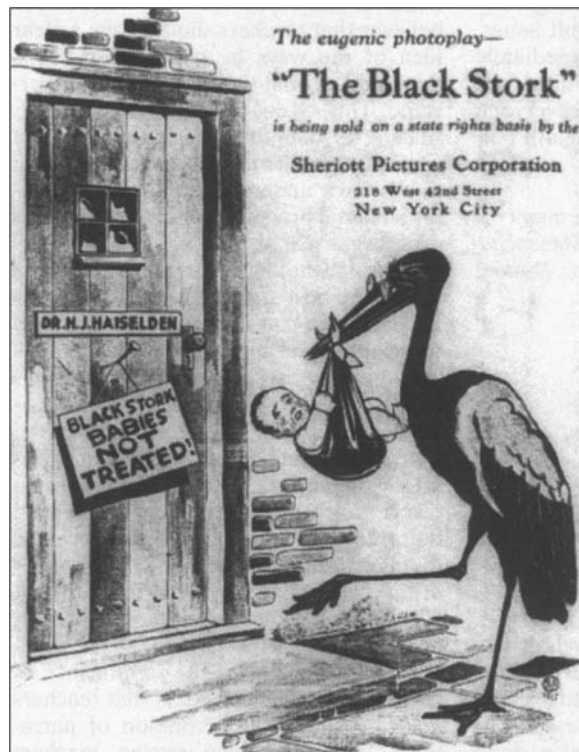
rather than written sources and in the mass media rather than professional texts. Attention to this dimension teaches us that the values endorsed by the eugenic leadership were not always shared by the public. A number of films ridiculed the belief that rational science should outweigh passionate love as the motive for mating. Eugenic leaders often encountered resistance in their efforts to shape public opinion. The public had a mind (or, rather, minds) of its own. Aesthetic responses were not necessarily those intended. Images of deformed bodies which were meant to evoke disgust often elicited fascination or pity instead.

Scientific readers of this book are likely

to be most interested in the insights they can glean for current biomedical debates. One lesson is that medicine can never be value-neutral. What counts as a disease or as a futile or humane intervention is not simply a technical question that can be answered by experts. Another lesson is that the distinction between allowing somebody to die and killing somebody is easily eroded. While Haiselden initially emphasized the difference between withholding treatment and active euthanasia, he later rejected the distinction. Apparently no one who supported his initial position explicitly repudiated his increasingly active methods. A third lesson is that humanitarian and cost rationales are readily blurred. As is the case today, considerations of social utility coexisted with discussions of what was best for the infant.

Haiselden and his allies argued that non-treatment would relieve individual suffering and eliminate a social burden. But relief from pain does not seem to have been a primary motive. Who knew if the newborns really suffered?

That many families suffered was clear. Haiselden received numerous letters of support from despairing parents, especially women, who felt he was the only person who really understood their situation. Pernick points out that the principle of autonomy in medical decision-making is thought to provide a barrier to the return of eugenics. We tend to attribute past abuses to an excess of power in the hands of the state. As long as individuals retain freedom of choice, we are confident that the use of medical technologies will be humane. Pernick reminds us that



At death's door — a 1917 advertisement for the film *The Black Stork*.

believed that "our puny sentimentalism has caused us to forget that a human life is sacred only when it may be of some use to itself and to the world", proposed establishing doctors' juries. She thought this method would produce objective decisions. But medical leaders feared that assuming decision-making authority would, in the end, weaken their profession. They also found Haiselden's publicity-seeking distasteful. Like Jack Kevorkian, Haiselden practised a "propaganda of the deed" that aimed to stimulate public debate and force people to take sides. But doctors thought the less said about the issue the better.

Pernick further suggests that the eugenics movement had a powerful aesthetic dimension, which has been overlooked because it appears most clearly in visual

eugenic euthanasia was chosen by families. It was a practice that resulted from individual rather than collective decisions.

I felt one frustration with *The Black Stork*. The author hopes that an analysis of their antecedents will shed light on the modern Baby Doe and right-to-die controversies. In many ways, it does. But while the history of the Bollinger case alerts us to issues that we might otherwise ignore, we are still left, in the end, needing to take a stand on the appropriateness of doctor-assisted suicide and parental non-treatment decisions. From the tone of the book, it appears that the author disapproves of both. If so, I wish he had said so directly. If not, I would like to know why. But at least *The Black Stork* will greatly enrich discussion of these difficult issues. While the book will most immediately attract readers interested in the history of eugenics and the history of film, everyone concerned with contemporary health policy will profit from reading it. □

Diane B. Paul is in the Department of Political Science, University of Massachusetts, 100 Morrissey Boulevard, Boston, Massachusetts 02125, USA.

A little learning

Peter Bryant

The Culture of Education. By Jerome Bruner. *Harvard University Press: 1996.* Pp. 224. \$24.95, £15.95.

JEROME Bruner is a famous psychologist, and justly so. He was one of the pioneers of cognitive psychology, the study of the ways in which people think and reason and solve problems, and he has also done much exciting research with children. The combination of these interests led him into research on education, a topic he wrote a great deal about in the 1960s and 1970s.

At that time, he was best known in educational circles for his idea of the spiral curriculum. He recommended that children should be taught at different ages about the same topic but at different levels of sophistication. Under this scheme they would be introduced to a topic in physics, for example, or in history, in a basic way reasonably early on in school, and then reintroduced to it in a more formal and abstract manner later.

Bruner later turned his energies to studying children's language and their interactions with their parents, and the way in which people form narratives. His interest in education appeared to wane. Now, however, in a spiral-like way he has revisited the subject after a long break, and has written a book reflecting further on how children should be taught.

The book has three main themes. One is the importance of the culture — by which Bruner means the social environment — in which schoolchildren live. He thinks that children's educational aspirations, their learning and their plans to use this learning in their adult life are all shaped by culture.

The second theme is what Bruner sometimes calls "intersubjectivity" and sometimes "theory of mind". He is concerned with the ability of humans to work out with a fair degree of accuracy what other people know or believe about a particular topic or event. He argues that this is an essential and desperately important thread in any form of education, but his thesis is sometimes hard to follow. He believes that teachers should have a clear idea of the ways in which their pupils think and of what they are likely to understand. This, of course, is not a new point, since it is one that Piaget made many years ago. But Bruner also claims that the pupils' own understanding of other people's mental processes plays a role in their progress at school, although it is not at all clear from what he writes why this should be so. The best argument I could find was that children can understand the nature of scientific or historical evidence only when they grasp the fact that people's beliefs, true or false, are based on evidence, which can be strong but may be weak. But Bruner does not develop this interesting theme, and tends instead to resort to phrases that sound impressive but are hard to understand whenever he discusses intersubjectivity. "The pedagogy of mutuality" has a certain ring to it, but it is difficult to see what is new about it apart from the term itself.

The final theme is the significance of narrative. Bruner considers that teachers have neglected the formation of narratives. In the arts, he argues, teachers accept that narrative is important, particularly for the study of history and literature; but, he claims, they do not teach it. "It has always been tacitly assumed that narrative skill comes 'naturally', that it does not have to be taught. But a closer look shows this not to be true at all." I find the claim that children are not taught about narratives hard to accept, partly because it seems not to be based on any evidence and partly because a considerable part of my own education and of my children's education involved instruction about how to write about historical or fictional sequences of events. Narratives are not always taken for granted at school.

Bruner's discussion of narratives and science is more promising. He argues that humans organize their knowledge in two main ways, one through logical-scientific thinking and the other through narrative thinking. The first kind of thinking forms the basis of science and mathematics teaching, whereas the second forms the

basis of teaching in the arts. Bruner deplores this divide and suggests that narratives can help scientific learning too. Narratives involve sequences of events and also an evaluation of these events, and so does scientific research; yet we do not present science to children in that way, and so rob them of the intrinsic excitement of the subject.

Here is an interesting start, but Bruner does not push his argument far and his discussion degenerates into an account of one item at an exploratorium and an anecdote about an apparently trivial conversation between the Queen of England and the historian Alan Bullock, which seems to be annoyingly beside what could be a really important point.

There are clear weaknesses, therefore, in two of Bruner's three main themes, and the book is also beset by a more general problem. For a work on education, it is alarmingly light on specific educational issues, such as how to teach children about algebra or reading or motion. This is a great pity because it is on questions like these that psychologists have made the most progress in recent years. Even when the book gets down to educational brass tacks, it does so far too cursorily. For example, Bruner briefly mentions the relationship between addition and multiplication as an example of deriving new knowledge from old, but he ignores all the fine work done in recent years on the connections and the disconnections between children's knowledge of addition and of subtraction.

Another mathematical example that he cites is the considerable amount of mathematics that adults use in their daily routines. Yet perversely he ignores the extraordinarily exciting and far more relevant research on children's informal mathematics which demonstrates that children who work in the informal economy outside school can learn sophisticated mathematical procedures in that context but are often quite incapable of using this valuable knowledge in mathematical classes at school. Another example of Bruner's neglect of good and relevant research is what he writes about children "going meta", to use his phrase. "Metacognition" means our (rather limited) ability to understand our own cognitive processes, and Bruner's thesis is that this form of cognition is important for children at school. But exactly how it helps is never explained in the book, which presents some anecdotes on children's metacognition but neglects the impressive body of empirical work on the topic, such as research into the relationships between metalinguistic abilities and learning to read.

The book is at its best in telling us about Bruner's own wide interests, for example in narrative, in the psychology of legal practices and in the nature of chil-

dren's theory of mind. If readers forget about education and concentrate instead on Bruner's views on broad topics such as these, they will be fascinated. If they want to know more about how to teach children, they will be disappointed. □

Peter Bryant is in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Collapse of culture

George M. Marsden

The University in Ruins. By Bill Readings. Harvard University Press: 1996. Pp. 238. \$29.95, £18.95.

WHEN so ancient an institution as New College, Oxford, attaches a statement of its dedication to "excellence" to all its job advertisements, it must be part of a cultural trend. So observes Bill Readings, who points out correctly that 'excellence' is the catch-all term used by university administrators when they really have nothing to say. "The general applicability of the term", he says, "is in direct relation to its emptiness." Yet the near universality of 'excellence' as the standard by which universities describe themselves also has ominous implications. It signals the demise of earlier meanings that justified the university's existence and the emptiness of the present alternative.

Readings has no nostalgia for earlier meanings for the university, but he is alarmed by the signs that universities are left with only a business management model that is appropriate to the globalization of capitalism. The justification for universities used to be that they were servants of the nation-state and promoters of its cultural ideals. Now they view themselves as business corporations. The nation-state (although not nationalism), says Readings, has declined as a definer of culture and increasingly is being replaced by the transnational business model. In the present model, questions of meaning have disappeared and questions of production have come to the fore. Achievement is measured quantitatively. It does not much matter what research is about just so long as measurable amounts continue to be produced. Teaching is increasingly gauged by quantitative standards. The comparative 'excellence' of departments in UK universities is measured by 'performance indicators' that help to decide government funding.

Modern universities, according to

Readings, went through two earlier stages that defined their overall meaning. The first is the version of modernity articulated by Immanuel Kant, which emphasized the centrality of reason. Philosophy was to be the integrating queen of the sciences and the universities would train rational citizens. In the nineteenth century this ideal was superseded by the triumph of 'culture' as articulated first by the German Idealists. Universities were justified by serving the nation-state, especially by cultivating appreciation of the nation's highest cultural achievements. In English-speaking countries, literature replaced philosophy as the pre-eminent culture-bearing discipline. Celebrations of the national literature and the Western canon served to imbue citizens with the highest ideals of the civilization.

Since the 1960s, this literary-cultural ideal has fallen apart as a justification for university education. Literature has been taken over by cultural studies. Readings, professor of comparative literature at the

insights, but it is difficult to know what counts as evidence for any particular explanation. Often, as is common in literary critical circles, points are supposedly settled by appeal to authority. That Derrida, Lyotard or Foucault made a sweeping assertion seems to count as making it, as it were, true.

The most glaring weakness in Readings' argument is that it says almost nothing of the roles of natural science and technology in the university. Readings explains briefly in the introduction that the sciences gained their powerful place in the university "by analogy with the humanities". C. P. Snow's 'two cultures' suggests that the sciences justify their roles in the university, especially their teaching role, on a cultural model. Because the book is about the collapse of 'culture' as a viable justifying narrative, the sciences also lose their higher meaning when the humanities deny any transcendent cultural norms.

Readings seems to take it for granted that natural science and technology are the ugly step-sisters in the piece. Both the nineteenth-century champions of 'culture' and the post-1960s destroyers of any grand metanarrative of 'culture' present their views as alternatives to technocratic-bureaucratic models that threaten to take over civilization. So Readings apparently sees little reason to discuss the sizeable roles of natural science and technology in shaping the university. Oddly, for instance, he says almost nothing of the way in which the research ideal of the natural sciences was applied to the humanities. Nor does he pay attention to the justification of modern universities as servants to the nation-state in developing science and technology.

The volume would be considerably better if it were entitled *The Humanities in Ruins*. On that subject it is often illuminating, even if lacking analytical rigour. The more modest title, moreover, does not preclude some wider implication. If the humanities are in ruins, it reflects a wider culture in ruins. As Readings suggests, our culture has little basis for establishing what we value other than what is currently popular in the market. A culture in ruins also implies universities in ruins. Readings is correct in saying that universities are as much businesses as is the National Basketball Association, and no one should have illusions that either stands for any higher ideals. Nevertheless, the substance of the book is all based on the self-destruction of the humanities and does not deal with the rest of the university in any direct way.

What can we do when we recognize that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

E. R. Jarrett/Architectural Association

New College, Oxford, seeks "excellence" — but does this signify the end of all it used to stand for?

University of Montreal before his untimely death in an air crash in 1994, focuses his study on this field. Sympathetic to 'left-wing' criticism, he nonetheless shows little patience with the pretensions of most of cultural studies. Contemporary cultural critics still present themselves as redeemers through culture. 'Multiculturalism' serves the purposes of the transnational corporations on which the university business is modelling itself. Cultural studies simultaneously promotes such a hegemonic cultural ideal and teaches that there is no basis for any such ideal.

Readings' historical analysis suffers from many of the typical traits of the literary analysts of culture he otherwise criticizes. Sweeping generalizations abound. The hegemony of the nation-state serves as the all-purpose explanation for anything that happened during the modern era and then is superseded by the nearly universal dominance of globalized capitalism. Pointing out this overall pattern leads to many worthwhile

the university has lost what was once supposed, at least by the humanists, to give it higher meaning or identity? Readings' own solution is quixotic, given his critical stance which seems to undercut everything by *ad hoc* temporary norms. He proposes that the university should be a "community of dissensus", where we can talk to and accept "the Other" while exercising no hegemonic control. He recognizes that a secondary consensus would be necessary to establish such a dissensus. He does not explain, however, what the advocates of dissensus or difference are to do with "the Other" who firmly hold to an alleged norm they regard as hopelessly benighted. Readings simply tells us to accept our dependency on others, even if we cannot say why. Such a "non-finite obligation" may make some people think about religion and the mystical, he observes, "but I am not trying to sound mystical". His moral ideal, which is roughly that of the Sermon on the Mount, is admirable. How contemporary cultural criticism could persuade modern university culture to adopt it remains a mystery. □

George M. Marsden is in the Department of History, University of Notre Dame, Notre Dame, Indiana 46556, USA.

Two for the skies

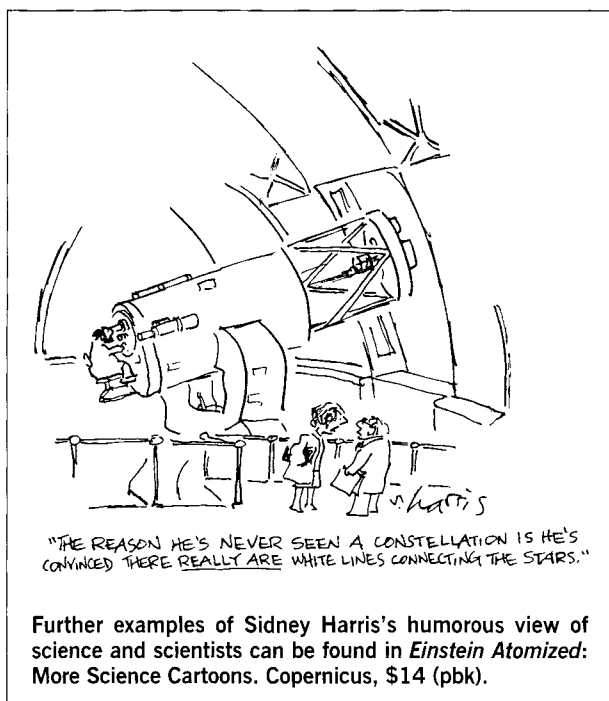
William H. Press

Blind Watchers of the Sky: The People and Ideas that Shaped Our View of the Universe. By Rocky Kolb. Addison-Wesley: 1996. Pp. 338. \$25.

Our Evolving Universe. By Malcolm S. Longair. Cambridge University Press: 1996. Pp. 185. £24.95, \$34.95.

Is basic science not getting the respect it deserves? Has 'curiosity' become a term of vilification in political quarters? While most of us in the trenches make slow progress through the pop-psych stages of denial, anger, grief and acceptance, a few of our colleagues are taking the more constructive and proactive approach of engaging the public through popular writing about basic, curiosity-driven science. Here are two books on modern cosmology, undeniably one of the purest of the basic sciences, which, in very different ways, succeed in conveying the vibrancy, creativity and sheer intellectual joy of 'this cosmology thing' and, by contagion, basic science in general.

Dr Jekyll, meet Professor Edward W. Kolb, respected author of some of the most arcane and inaccessible papers in the specialized field of particle astrophysics. Mr Hyde, meet Rocky Kolb, gutsy narrator of a thoroughly delightful recounting of the



great ideas of the past 500 years of cosmological understanding, and no less the very human personalities who have given them to us. Use of his nickname, Kolb tells us, "was to serve as a constant reminder that this book is not for the usual target audience for whom I write".

The play seems to have succeeded. No less successful is the daring strategy of prefacing a book on modern cosmology with 130 pages on Copernicus, Tycho Brahe, Kepler, Galileo and Newton. In the hands of 'Edward W.' this might have been dry history, but Rocky delivers (on the basis of secondary sources, but quoting liberally from the principals) a fascinating narrative, full of audacious wit and genuinely new insight, on what might have otherwise been assumed to be a thoroughly explored historical trail.

What interests Kolb, apart from tracing the history of ideas (which he does well), are questions such as "How did the personal backgrounds and eccentricities of the 'greats' affect their discoveries?" and "What parts of the modern scientific paradigm did they 'get', and which parts were they blind to?"

Kolb has a novelist's knack for projecting us backwards into epochs very different from now, and his analysis, given its generally high entertainment value, is surprisingly sophisticated. Galileo, in his view, is the only one of the bunch who could really fit right into a modern university department. (Now there's a recruitment opportunity that would make any dean drool!)

Jumping forward in time, Kolb takes up in similar vein the discovery of the Galaxy, and of the now accepted Big Bang model. A number of lesser figures have cameo roles here (Messier, the Herschels, Kirchoff, Hertzsprung and Russell) but

the detailed narrative focuses, as it should, on Shapley (who got it wrong), Einstein, Friedmann and Hubble (who got it right). No less interesting are such supporting players as Humason, who, without benefit of formal education beyond elementary school, rose from mule driver and construction worker to become Hubble's essential collaborator.

In the final section, Kolb introduces the cosmic microwave background and provides an account of the state of current understanding of the early moments of the hot Big Bang. Clear discussion of the science now overshadows the personalities and egos involved. One can hardly

believe that the author's gossipy interest in people (and what causes them to do good or bad science) ends with the year 1965. Nor, I can attest, have all modern cosmologists become self-effacing specimens of the dull bourgeoisie. Can it be that Rocky is less willing to take on living targets? If so, we are deprived of some undoubtedly witty, thoughtful and provocative commentary. Perhaps another book, under a different, more clandestine, pseudonym, is in order.

Malcolm Longair's *Our Evolving Universe* is a more conventional, though certainly no less authoritative, introduction to modern cosmology. The author, formerly Astronomer Royal for Scotland and now a professor at the University of Cambridge, has given us a beautifully produced book, in an oversize format, with an elegant typographical design, clear diagrams and many colour illustrations. He begins by discussing some observational aspects of astronomy, then turns to the formation and evolution of stars, the origin of quasars, then galaxies and (finally) the Universe in its entirety. The text is unflinchingly direct and accurate; although there are no equations, the discussion is at a uniformly high conceptual level.

I hope that there is a market for a book like this. Its content is that of a first-rate introductory course for non-astronomy majors, while its clean format, with none of the intrusive pedagogical devices that make textbooks *qua* textbooks so annoying to the general reader, makes it a delight to hold. □

William H. Press is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

Seismological evidence for differential rotation of the Earth's inner core

Xiaodong Song & Paul G. Richards

Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA, and Department of Earth and Environmental Sciences, Columbia University, New York, New York 10027, USA

The travel times of seismic waves that traverse the Earth's inner core show a small but systematic variation over the past three decades. This variation is best explained by a rotation of the inner core that moves the symmetry axis of its known seismic anisotropy. The inferred rotation rate is on the order of 1° per year faster than the daily rotation of the mantle and crust.

THE Earth's liquid iron core was formed very early in the planet's history, after which slow cooling has led to the formation and growth of the solid inner core. The inner core was discovered by seismological methods in 1936¹. Its slow growth to a present radius of about 1,220 km has provided a source of energy as liquid iron freezes at the inner-core boundary. This energy is thought to drive convection in the outer core, which in turn maintains the dynamo action that generates the Earth's magnetic field².

As the inner core lies at the centre of a much larger, liquid outer core (radius 3,480 km) of very low viscosity^{3,4}, it is easy for the inner core to rotate^{5,6}. We present observational evidence for inner-core rotation, at a rate that has taken the inner core through more than a quarter of a complete revolution this century, in an eastward direction. Around the equator of the inner core this rate corresponds to a speed of a few tens of kilometres per year, which is more than 100,000 times faster than the fastest relative motion of the tectonic plates of the lithosphere. Our estimate of the inner-core rotation rate comes from an interpretation of measurements of the difference in travel time between seismic body waves that pass through the inner core, and those that pass only through the outer core. We show that these measurements, made upon seismograms in archives going back to the mid-1960s, have systematically changed with time.

Essential to our interpretation of the travel-time data is the recent discovery that the inner core is anisotropic, with cylindrical symmetry about an axis aligned approximately with the Earth's north-south spin axis. Such anisotropy was first proposed^{7,8} to explain the faster propagation of seismic waves that travel closer to the spin axis⁹ and the anomalous splitting of certain normal modes of the whole Earth that have significant energy in the inner core¹⁰, and has been recently confirmed^{11–13} to have a significant amplitude of about 3%. The fact that the inner core is anisotropic is now well established, based on further studies of travel times and normal modes^{14–22}. The cause of the inner-core anisotropy is believed to be preferred orientation of hexagonal close-packed iron (a high-pressure polymorph of iron²³), but how the alignment occurs is still under debate^{24–26}. Recent studies^{11,15,18} also indicate that the inner core's axis of anisotropy is tilted by a few degrees from the north-south axis. A rotation about any axis different from the symmetry axis, including the north-south axis, will cause a systematic change in the times of waves transmitted through the inner core from a repeating seismic source, observed at the same fixed station. Moreover, we can use the arrival times of seismic waves transmitted through the outer core as reference points, as the outer core is thought to have negligible fluctuations in seismic velocity associated with its own convection²⁷.

We were led to our search for inner-core motion in part by the work of Glatzmaier and Roberts^{28,29}, who recently reported a three-dimensional, self-consistent numerical model of the geo-

dynamo. Their model has an inner core with a differential rotation that is "as large as a few times 10^{-9} rad s⁻¹ and changes on a timescale of about 500 years"²⁹ and is almost always rotating slightly faster than the mantle and crust. The rate that we report here is comparable with their value, is large enough to be important in many areas of geophysics and geochemistry, and adds a new dimension—time—to traditional seismological studies of the Earth's deep interior structure.

Differential travel-time observations

All of the seismic waves discussed here are simple short-period longitudinal waves (sound waves) with a wavelength of 5–20 km. Such waves that have propagated through the outer core are generically labelled PKP (Fig. 1). There are two possible paths, called PKP(AB) and PKP(BC), for waves that turn at the middle and the bottom of the outer core, respectively. The wave transmitted through both the outer and inner core is called PKP(DF). Our basic measurement is the time difference between the PKP(BC) and PKP(DF) arrivals, often abbreviated as the BC – DF time. Figure 1 shows that these two waves travel very closely together in the crust and mantle and in most of the outer core, so that the differential time (BC – DF) is relatively insensitive to uncertainty in source location and to three-dimensional heterogeneities of the crust and mantle along the ray paths. A standard Earth model (the Preliminary Earth Reference Model, or PREM³⁰) is used to form the BC – DF residual (that is, the observed time difference minus the time difference predicted for PREM) for purposes of subsequent analysis. In addition, residuals of the differential times between PKP(AB) and PKP(BC), which pass through the outer core only, are used to examine possible systematic biases in source locations. Residuals of the differential times between PKP(AB) and PKP(DF) are also useful when PKP(BC) is not observable.

We began to explore time-variability in PKP(DF) times using north-south paths from nuclear explosions at Novaya Zemlya

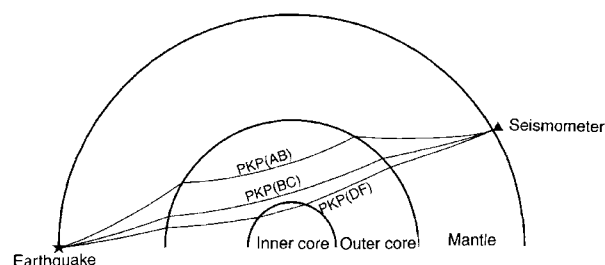


FIG. 1 Ray paths for seismic waves that pass through the Earth's cores.

(which have well constrained locations at around 73° N, 54° E) to seismographic stations in Antarctica. Direct visual comparisons of AB – DF times in the analogue seismograms from different explosions show a decrease by about 0.2 s after about a decade—a difference not much greater than the resolution of the direct visual comparison. To improve the precision of measurement significantly will require scanning and digitization of the analogue data, ideally using the original paper records instead of microfilms (an effort that we have initiated).

We therefore turned to stations with significantly long histories of digital recordings, which permit even more precise measurement of travel times than digitized analogue data. One such station is College, Alaska (station code COL), which has 14 years of digital recordings to date (1982–96), in addition to 18 years of World Wide Standardized Seismographic Network (WWSSN) analogue recordings (1964–82). In our subsequent analysis, the data for the 1960s and 1970s are digitized from the corresponding analogue records. The north–south paths from earthquakes in the South Sandwich Islands (between South America and Antarctica) to COL had provided critical observations in previous anisotropy studies^{11,12}. Figure 2 shows an overlay of PKP waveforms from such paths for two earthquakes that occurred about 15 years apart in almost exactly the same location. The time differences between BC and AB are almost exactly the same, confirming that the two event locations are indeed very close. The DF signal for the 1982 record, however, arrives earlier by about 0.4 s relative to BC than the DF signal for the 1967 record.

We use records such as those shown in Fig. 2 to measure BC – DF and AB – BC differential times by cross-correlation techniques. The BC – DF and AB – BC residuals obtained for the South Sandwich Islands to COL paths for a period of about 28 years (1967–95) are divided into three groups, about 12 years apart (Fig. 3). It can be seen that the BC – DF residuals increase from the 1967–75 period, to the 1980–85 period, and to the 1992–95 period (Fig. 3a), the increase amounting to about 0.3 s over the 28-year period. Possible explanations include a change in the DF (or BC) travel times, or an effect from systematic mislocation of the events. An increase in source–station distance by only 50 km would change the BC – DF time at distances near 151° by 0.3 s, so the time change could potentially be an artefact of the use of different global networks, used to locate the earthquakes in different decades. However, the distribution of the AB – BC residuals (Fig. 3b) for those records which have BC – DF measurements does not appear to have a systematic change with time, suggesting that the observed variation in AB – DF residuals is unlikely to be due to systematic event mislocations or recording instrument biases.

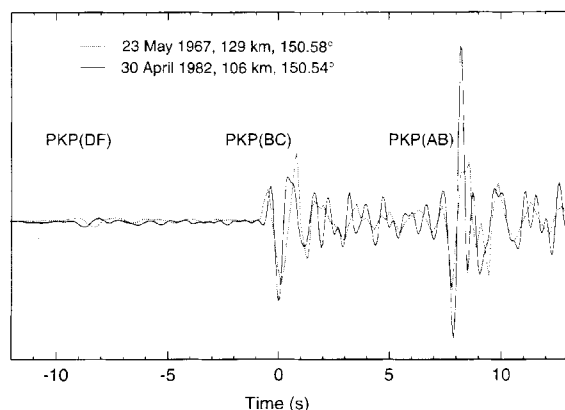


FIG. 2 An overlay of two short-period PKP seismograms from earthquakes that occurred 15 years apart at almost the same location in South Sandwich Islands, as recorded at College, Alaska (COL). Event locations in this study are taken from the Earthquake Data Reports of the US Geological Survey.

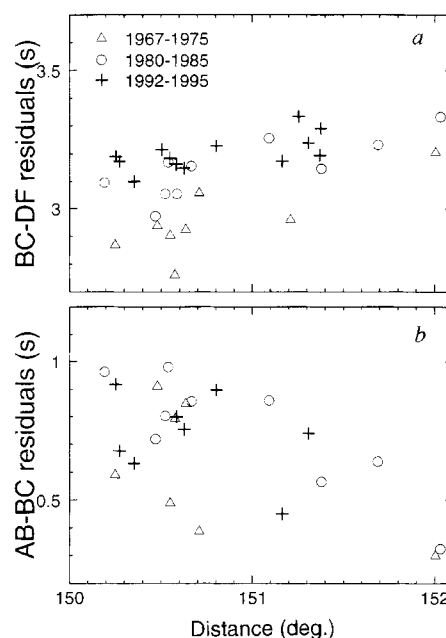


FIG. 3 Residuals for BC – DF (panel a) and AB – BC (panel b) times obtained for earthquakes grouped in three time intervals in the same region (South Sandwich Islands), observed on north–south paths at the same station (COL). Distance is measured by the angle subtended at the Earth's centre by the earthquake source and seismometer locations. The original broadband data for the 1990s are converted to the same short-period instrument response for the earlier years before the measurements. The BC – DF residuals typically increase with time. No such pattern is apparent for the AB – BC residuals, indicating that systematic event mislocations are unlikely to be the cause of time-dependent effects on BC – DF.

Figure 4 shows BC – DF residuals for three event–station pairs that are chosen to have ray directions in the inner core that make very different angles to the equatorial plane. The residuals for different distances and focal depths are corrected to a standard distance and depth by dividing the observed BC – DF residual by the time that the DF ray spends in the inner core for the distance and focal depth concerned, and multiplying by the time the DF ray spends in the inner core for a source at 151° and depth of 0 km. The residuals for South Sandwich Islands to COL paths are progressively larger from the 1960s/1970s to the 1990s, as seen before. But the paths from Kermadec to Kongsberg (KONO) and Bergen (BER), Norway, show smaller residuals for the 1990s than the 1980s. The paths from Tonga to Graefenberg, Germany (GRFO) show no such systematic differences in the data between the 1980s and 1990s. The ray directions in the inner core for the Tonga–GRFO paths are so much closer to the equatorial plane that no effect of inner core anisotropy would be expected. The deviation of the observations at COL and KONO/BER from the time-independent model (dashed curve in Fig. 4) reflects a departure of the observations from a simple constant cylindrical anisotropy of the inner core with the symmetry axis parallel to the north–south axis. We conclude from Figs 3 and 4 that there is the need to consider the effects of more complex anisotropic structure, including a time dependence.

Figure 5 summarizes our observations at COL as a function of the calendar time at which the South Sandwich earthquakes occurred—and also interprets them with a time-dependent model. We observe a gradual increase in the residuals from early 1967 to early 1995, suggesting that the DF ray during this period was taking a progressively faster path through the inner core. Such faster paths can be achieved by a mechanism that decreases the angle between the path within the inner core and the axis of anisotropic symmetry. As the ray path is essentially fixed, we propose that the symmetry axis is moving, owing to a rotation

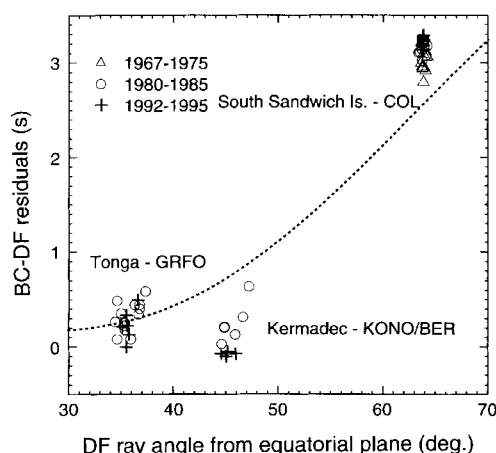


FIG. 4 Residuals for BC – DF times as a function of the angle between the inner-core part of the ray path and the equatorial plane. Ray paths to three different seismographic stations are used (COL—College, Alaska; GRFO—Graefenberg, Germany; KONO/BER—Kongsberg and Bergen, Norway), each using a separate set of earthquakes in the same source region (Tonga to GRFO; Kermadec to KONO/BER; South Sandwich Islands to COL). The dashed line is the prediction for the anisotropy model of Song and Helmberger¹², which has the axis of symmetry of the inner core aligned with the north–south axis. Systematic shifts in residuals over time are absent from the near-equatorial path, and are in opposite directions for the other two paths. Such systematic shifts indicate the need to consider time-dependent effects on the fixed paths through the inner core.

of the inner core. The dashed line in Fig. 5 shows our best estimate of the inner-core rotation required to explain the observed change of residuals over time, as discussed below. In this estimate, the inner core rotates about the north–south axis like the rest of the planet, but at a rate that is 1.1° per year faster. Because the symmetry axis of the anisotropy is not aligned north–south, it moves to the east as illustrated in Fig. 6.

Inferred inner-core motions

We have presented evidence that the inner core is moving, presumably by some type of rotation. But with the limited number of observations made so far, we can get only an approximate estimate of the rotation rate and orientation.

The inner core's moment of inertia, although only 0.07% of the Earth's total moment of inertia, is about 500 times larger than the moment of inertia of the atmosphere, whose internal motions can effect small changes in the Earth's length of day³¹. Observed changes in length of day, over a timescale of years, have been adequately explained without postulating transfer of angular momentum to or from the inner core. We therefore assume that the rotation axis and the rotation rate of the inner core do not change on the timescale of our observations.

At least eight parameters are needed to describe an appropriate model of the time variations in travel-time residuals: the latitude and longitude of the axis of symmetry at a particular time, a direction-dependent relative velocity perturbation from PREM in the form $a + b \cos^2 \xi + c \cos^4 \xi$ (where ξ is the angle between the ray direction in the inner core and the anisotropy symmetry axis), the latitude and longitude of the fixed rotation axis, and finally the fixed rotation rate.

We carried out a grid search for the latitude and longitude of the anisotropy symmetry axis to obtain values of a , b and c and that best fitted observed travel-time residuals without rotation. Using the BC – DF measurements of Song and Helmberger¹² and the new measurements of this study with a combined time span of the past 30 years, the time-invariant best fit to the anisotropy is given by

$$\Delta v/v = 0.0042 - 0.0256 \cos^2 \xi + 0.0517 \cos^4 \xi \quad (1)$$

with symmetry axis crossing the Northern Hemisphere at 82.0° N,

175.0° E. This best fit achieves a variance reduction from the original measurement (with respect to the isotropic PREM) of 93%, not significantly better than the variance reduction of 89% achieved without allowing for a tilt of the symmetry axis. Nevertheless, our estimate of the latitude and longitude of the axis is remarkably close to previous best-fitting symmetry axes. It is only 3.5° away from 79.5° N, 160° E of Su and Dziewonski¹⁸, estimated using PKP(DF) times of the International Seismological Centre Bulletins for years 1964–90; and 8.5° away from 80° N, 120° E of McSweeney *et al.*¹⁹, estimated using differential PKP times for events in the 1980s and 1990s, as in this study. Note the range of these latitude estimates is less than 3° .

The BC – DF measurements can be used to place a minimum bound on the rotation rate, even when the orientation of the symmetry axis of anisotropy is unknown. We computed this bound by calculating the rotation rate needed to match the 0.31-s BC – DF time difference over 28 years at COL (that is, the trend of the data in Fig. 5) for every possible orientation in a grid search, and then selected the smallest. The parameters of anisotropy obtained via the time-invariant analysis were used for this calculation. When the orientation of the rotation axis is also considered unknown, the minimum bound is 0.15° per year. The minimum bound increases if we assume that the differential rotation of the inner core is around the north–south axis. With this rotation axis, the minimum rate increases as the tilt of the symmetry axis decreases, from 0.45° per year for a 15° tilt to 1.30° per year for a 5° tilt. If the BC – DF time differences at KONO are used, the minimum rates are even higher. However, the trend is less well constrained owing to fewer samples.

If we limit the tilt of the symmetry axis to between 8° and 11° away from the north–south axis, as indicated by the above time-invariant analysis, but allow the longitudinal location to vary, the least-squares solution for rotation about the north–south axis to fit the time-varying residuals at COL and KONO puts the symmetry axis at 79.2° N, 169° E at the beginning of 1996 with a rotation rate of 1.1° per year. We regard this as our best estimate, and it gives a variance reduction for COL data compared to the best-fitting time-invariant model (that is without inner-core rotation) amounting to 64%. Assuming the tilt of the symmetry axis in the best-fitting model, the allowed minimum and maximum rotation rates are estimated to be 0.4° per year and 1.8° per year, respectively, from the minimum and maximum slopes of the COL residuals (Fig. 5) constrained by the two-standard-deviation error bars. The variance reduction for the KONO data is insignificant but note that they cover only 13 years (instead of 28 years at COL), and the earlier data are quite scattered. Nevertheless, our pre-

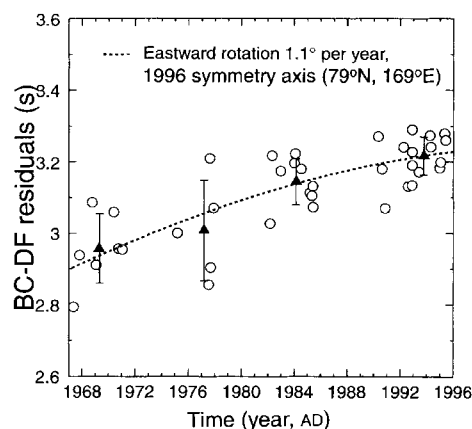


FIG. 5 Residuals for BC – DF times for South Sandwich Islands to COL paths as a function of earthquake time of occurrence. The observed residuals are indicated by open circles. The mean (filled triangles) and $\pm$ one standard deviation (error bar) are shown for the residuals over about four-year periods. The predicted times (dashed curve) are based on our best estimate of inner-core rotation about a north–south axis.

ferred model (Fig. 6) of inner-core motion has its symmetry axis moving away from the Kermadec–Norway paths, consistent with the observed decrease in BC – DF residuals for these paths (Fig. 4). The fact that the preferred model does not significantly improve the fit to the KONO data suggests a need for more data samples gathered for the Kermadec–Norway paths and other paths that are sensitive to the motion of the anisotropic symmetry axis.

The main assumption underlying our best estimate of the rotation rate, namely that the rotation axis is north–south, will be examined when additional data are brought to bear. However, we can already be fairly sure from seismological observations that the rotation axis is not likely to be near the Equator, as such a rotation would cause highly nonlinear changes of travel times for north–south ray paths when the symmetry axis is rotated towards the Equator. Such changes are not seen in the shifts of travel times at COL in Fig. 5. Also in support of an axis far from the Equator is the fact that going further back, to the mid-1960s, it is still true (as today) that only paths close to the north–south spin axis show fast anomalies^{12,16,18,21}.

Implications of detectable inner-core motions

Much tighter constraints on the rate of inner-core rotation and on the orientation of the rotation axes and symmetry axes will be possible as more recordings of north–south PKP waves are gathered and better event locations are obtained (through, for example, joint hypocentre determination^{32,33}, a simultaneous calculation of all hypocentres for a group of earthquakes, which gives better relative event locations). Systematic shifts in PKP(DF) times are best observed where there is the greatest change in $\Delta v/v$ with respect to ξ in equation (1), that is, $10^\circ \leq \xi \leq 45^\circ$. Also, the effect of inner-core rotation is likely to be amplified for waves that traverse the innermost inner core because the paths within the anisotropy are longer. Travel-time anomalies in such waves, twice as big as the anomalies for South Sandwich Islands to COL paths, have been observed^{17,21}. In addition to the differential travel-time data, it is possible that the absolute travel-time data of Su and Dziewonski¹⁸ and the normal-mode analyses of Tromp^{13,22} can be reinterpreted to include possible inner-core motions. Techniques based on reflected waves may also be applicable, if lateral variability of the inner core boundary resulting in variable reflectance can be found, analogous to medical uses of ultrasound reflections to monitor organs that move inside the human body.

Numerous projects are suggested by our conclusion that inner-core motion is observable and can be quantified. New estimates of viscosity in the outer core may become available that in turn may provide information on melting conditions and temperature through the outer core, and refined estimates of the heat flux across the core–mantle boundary. New constraints may be placed on the dynamo motions underlying the geomagnetic field, leading to refinement of models that can reproduce secular variation of the field and geomagnetic reversals. New questions also arise regarding the distribution of angular momentum throughout the



FIG. 6 A representation of our conclusion that the anisotropic symmetry axis of the inner core is moving owing to inner-core rotation. As the solid core and its symmetry axis rotates, there is a change in the angle between the symmetry axis and the inner-core leg of a ray for a fixed source and fixed station, resulting in the systematic shift that we see in the total time taken to travel along the ray. We show the motion of the symmetry axis since AD 1900 according to our best estimate of a 1.1° per year eastward rotation about the north–south axis.

Earth. For example, does the inner-core rotation axis track the north–south spin axis in the precession of the equinoxes? Does the rotation of the inner core result in observable changes in the gravity field?

The organization and maintenance of seismogram archives will need to be modified in order to facilitate studies of core motions. Older data should no longer be superseded by newer, but instead are vital to the accurate determination of rates of change. In this regard it is unfortunate that important sets of digital and analogue data—some of the best in the world, for our purposes—have deteriorated and/or have been destroyed. Support may be needed to recover data from certain archives that have fallen into disuse, or which were not designed for use by the research community.

Gubbins *et al.*³⁴ suggested that “All discussion of modes of core convection will remain speculative until some direct or indirect observational evidence is found.” We believe that observational evidence is now available. The Earth’s dynamo is in part composed of a rotating inner core. □

Received 15 March; accepted 25 June 1996.

- Lehmann, I. *Publ. Bur. Cent. Assoc. Int. Seismol.* **A14**, 87–115 (1936).
- Braginskii, S. I. *Dokl. Akad. Nauk SSSR* **149**, 8–10 (1963).
- Gans, R. F. *J. geophys. Res.* **77**, 360–366 (1972).
- Poirier, J. P. *Geophys. J.* **92**, 99–105 (1988).
- Steenbeck, M. & Helmig, G. *Geophys. J. R. astr. Soc.* **41**, 237–244 (1975).
- Gubbins, D. *J. geophys. Res.* **86**, 11695–11699 (1981).
- Morelli, A., Dziewonski, A. M. & Woodhouse, J. H. *Geophys. Res. Lett.* **13**, 1545–1548 (1986).
- Woodhouse, J. H., Giardini, D. & Li, X.-D. *Geophys. Res. Lett.* **13**, 1549–1552 (1986).
- Poupinet, G., Piliet, R. & Souriau, A. *Nature* **305**, 204–206 (1983).
- Masters, G. & Gilbert, F. *Geophys. Res. Lett.* **8**, 569–571 (1981).
- Creager, K. C. *Nature* **356**, 309–314 (1992).
- Song, X. D. & Helmberger, D. V. *Geophys. Res. Lett.* **20**, 2591–2594 (1993).
- Tromp, J. *Nature* **366**, 678–681 (1993).
- Shearer, P. M., Toy, K. M. & Orcutt, J. A. *Nature* **333**, 228–232 (1988).
- Shearer, P. M. & Toy, K. M. *J. geophys. Res.* **96**, 2233–2247 (1991).
- Shearer, P. M. *J. geophys. Res.* **99**, 19647–19659 (1994).
- Vinnik, L., Romanowicz, B. & Breger, L. *Geophys. Res. Lett.* **21**, 1671–1674 (1994).
- Su, W. J. & Dziewonski, A. M. *J. geophys. Res.* **100**, 9831–9852 (1995).
- McSweeney, T. J., Creager, K. C. & Merrill, R. T. *Phys. Earth planet. Inter.* (submitted).
- Song, X. D. & Helmberger, D. V. *J. geophys. Res.* **100**, 9805–9816 (1995).

- Song, X. D. *J. geophys. Res.* (in the press).
- Tromp, J. *GSA Today* **5**(7), 137–151 (1995).
- Takahashi, T. & Bassett, W. A. *Science* **145**, 483–486 (1964).
- Jeanloz, R. & Wenk, H. R. *Geophys. Res. Lett.* **15**, 72–75 (1988).
- Karato, S. *Science* **262**, 1708–1711 (1993).
- Stixrude, L. & Cohen, R. E. *Science* **267**, 1972–1975 (1995).
- Stevenson, D. J. *Geophys. J. R. astr. Soc.* **88**, 311–319 (1987).
- Glatzmaier, G. A. & Roberts, P. H. *Nature* **377**, 203–209 (1995).
- Glatzmaier, G. A. & Roberts, P. H. *Phys. Earth planet. Inter.* **91**, 63–75 (1995).
- Dziewonski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297–356 (1981).
- Lambeck, K. *The Earth’s Variable Rotation: Geophysical Causes and Consequences* 275–282 (Cambridge Univ. Press, London, 1980).
- Douglas, A. *Nature* **215**, 47–48 (1967).
- Dewey, J. W. & Algermissen, S. T. *Bull. seism. Soc. Am.* **64**, 1033–1048 (1974).
- Gubbins, D. C., Thomson, C. J. & Whaler, K. A. *Geophys. J. R. astr. Soc.* **68**, 241–251 (1982).

ACKNOWLEDGEMENTS. We thank the IRIS DMC for easy access to data; the US NSF for long-term funding of the WWSSN archive at Lamont; J. Dewey and R. Engdahl for help with event relocations; and W. Menke, D. Kent and A. Lerner-Lam for comments on early drafts. This work was supported by the US NSF.

CORRESPONDENCE should be addressed to X.S. (e-mail: xsong@LDEO.columbia.edu).

A new member of the *frizzled* family from *Drosophila* functions as a Wingless receptor

Purnima Bhanot*, Marcel Brink[†], Cindy Harryman Samos[†], Jen-Chih Hsieh^{*‡}, Yanshu Wang^{*‡}, Jennifer P. Macke^{*‡}, Deborah Andrew[§], Jeremy Nathans^{*‡||} & Roel Nusse[†]

* Department of Molecular Biology and Genetics, § Department of Cell Biology and Anatomy, || Department of Neuroscience, and ‡ Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Room 805 PCTB, Baltimore, Maryland 21205, USA

† Howard Hughes Medical Institute and Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94305, USA

Receptors for Wingless and other signalling molecules of the *Wnt* gene family have yet to be identified. We show here that cultured *Drosophila* cells transfected with a novel member of the *frizzled* gene family in *Drosophila*, *Dfz2*, respond to added Wingless protein by elevating the level of the Armadillo protein. Moreover, Wingless binds to *Drosophila* or human cells expressing *Dfz2*. These data demonstrate that *Dfz2* functions as a Wingless receptor, and they imply, in general, that Frizzled proteins are receptors for the *Wnt* signalling molecules.

There is abundant evidence that secreted Wnt proteins have important signalling functions during animal development. For example, Wnt proteins have been implicated in cell-lineage decisions in *Caenorhabditis elegans*, in embryonic and adult pattern formation in *Drosophila*, in axis formation and dorsal–ventral polarity determination in *Xenopus* embryos, and in central nervous system (CNS) development and oncogenesis in mice^{1–3}. However, the Wnt proteins have been difficult to obtain in a soluble form, a problem that has hampered the development of biochemical and cell biological assays. Most information about the mechanism of Wnt signalling has come instead from the genetic analysis of *Drosophila* segment polarity and the role of the *Wnt* gene *wingless* (*wg*; refs 2,4–8). Within each embryonic segment, production of the *wingless* protein (Wg) by a narrow stripe of cells maintains *engrailed* expression in an adjacent stripe of cells.

In the embryonic epidermis the *wg* signalling pathway is defined by several genes: *dishevelled* (*dsh*)^{9,10}, *zeste white 3* (*zw3* or *shaggy*); and *armadillo* (*arm*), a member of the beta-catenin gene family¹¹, which is thought to be inactivated by *zw3*. The *wg* signal seems to counteract the inhibitory effect of *zw3*, leading to activation of *arm*^{12–14}. In *Drosophila* embryos the cytoplasmic levels of the *arm* protein (Arm) are increased as a consequence of *wg* signalling¹⁵. As judged by sequence data, none of the proteins identified thus far in the signalling pathway is a Wg receptor.

On the basis of genetic interactions between *wg* and other genes in the *wg* pathway, we have established a tissue-culture system for *wg* signalling¹⁶. In this assay, Wg produced by *Drosophila* S2 cells is added in soluble form to a cell line (clone 8) derived from *Drosophila* imaginal discs¹⁷. Like *Drosophila* embryos, clone 8 cells respond to Wg by specifically increasing the levels of hypophosphorylated Arm¹⁶, suggesting that these cells express a receptor specific for Wg.

Here we report the identification of a novel *Drosophila* gene, *frizzled2* (*Dfz2*), and demonstrate that it functions as a Wg receptor in cultured cells. *Dfz2* was identified in the course of characterizing a large family of vertebrate and invertebrate homologues of the *Drosophila* gene *frizzled* (*fz*)¹⁸. Mutations in *fz* result in aberrant orientations of adult cuticular structures, a tissue polarity phenotype^{19–21}. The *fz* sequence predicts an encoded protein with an amino-terminal cysteine-rich extracellular

domain followed by seven transmembrane segments^{22,23}. These characteristics have led to the suggestion that *fz* is a receptor for an unidentified ligand that transmits tissue-polarity information²⁴. Although *wg* does not seem to be involved in the tissue-polarity pathway and *fz* does not seem to be involved in the segment-polarity pathway, a possible link between the two pathways is suggested by the requirement in each for *dsh* function^{10,25}.

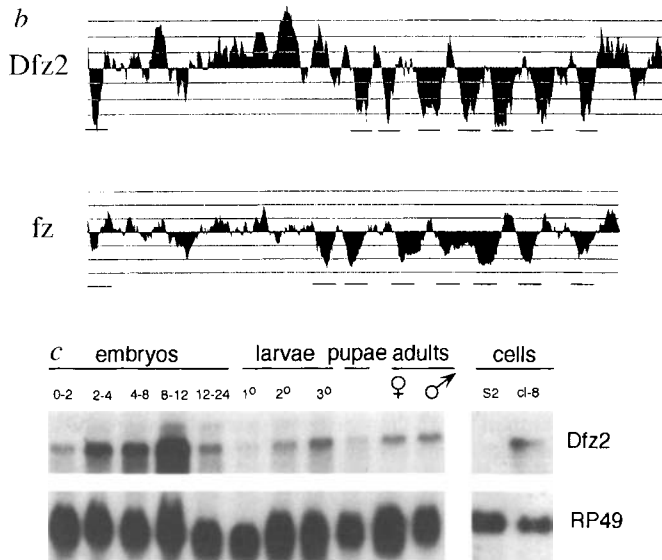
Molecular cloning of *Dfz2*

Using the sequences of *fz* and three mammalian *fz* homologues, degenerate polymerase chain reaction (PCR) primers were designed for the purpose of amplifying additional *fz*-like sequences¹⁸. PCR amplification using *Drosophila* genomic DNA as a target revealed a novel *frizzled* family member, *Drosophila frizzled2* (*Dfz2*). Isolation and sequence analysis of genomic and complementary DNA clones corresponding to *Dfz2* revealed a single coding exon containing an open reading frame of 694 amino acids (Fig. 1a). The predicted *Dfz2* protein (*Dfz2*) resembles all other members of the *frizzled* family in having the following structural motifs (beginning at the N terminus): a putative signal sequence, a domain of 120 amino acids with an invariant pattern of ten cysteine residues, a highly divergent region of 40–100 largely hydrophilic amino acids that is predicted to be flexible, and seven putative transmembrane segments (Fig. 1b). The C terminus of *Dfz2* resembles that of most mammalian *frizzled* protein in ending with the sequence S/T-X-V. A comparison with all known *frizzled* sequences shows that *Dfz2* most closely resembles human *fz5* and mouse *fz8* with which it shares 49% and 45% amino acid identity, respectively. *Fz* and *Dfz2* share 33% amino-acid identity. The *Dfz2* gene resides at 76A on the polytene chromosome map as determined by *in situ* hybridization (data not shown).

Developmental expression of *Dfz2*

As a first step in elucidating the function of *Dfz2* we examined temporal and spatial expression patterns by northern blot analysis, *in situ* hybridization, and immunostaining. A 5.5-kilobase (kb) *Dfz2* transcript is found throughout the *Drosophila* life cycle, most prominently during embryogenesis and in late larval and pupal life (Fig. 1c). At 2 hours post-fertilization, embryos have low levels of *Dfz2* RNA, which is presumably of maternal origin. *Dfz2* expres-

FIG. 1 Isolation and characterization of *Dfz2*—a novel *frizzled* family member in *Drosophila*. **a**, Alignment of the predicted amino-acid sequences of *Dfz2* and *Fz*. Identical residues are indicated in the consensus and conserved cysteines in the cysteine-rich domain¹⁸ are in bold face. **b**, Hydropathy profiles of *Dfz2* and *Fz* proteins were calculated according to ref. 37 with a window size of 15 amino acids; increasing hydrophobicity is downwards. The seven putative transmembrane domains and the signal peptide are indicated by a solid line. **c**, Expression of *Dfz2* RNA during *Drosophila* development and in clone 8 and S2 cells. Upper panel, a 950-bp segment from the *Dfz2* coding region was used to probe a blot containing 40 µg RNA per track from the indicated stages of *Drosophila* development or 30 µg total RNA from S2 cells and clone 8 cells. For the embryo samples, times during development are indicated in hours. A 5.5-kb transcript was detected at all stages, including 0–2 h post-fertilization. Lower panel, a ribosomal protein 49 (RP49) probe was used to probe the same blot to control for sample loading and transfer efficiency. **METHODS.** *D. melanogaster* genomic DNA was used as a template for PCR amplification using fully degenerate primers corresponding to conserved sequences YPERPI (sense) and WFLAA (antisense) as described¹⁸. A single PCR product corresponding to codons 348–425 of *Dfz2* was obtained and used to screen a *D. melanogaster* genomic DNA library in bacteriophage λ. Sequence analysis of cloned genomic DNA revealed an open reading frame of 694 codons with 33% amino-acid identity to *Fz*. Partial sequences from 6 independent cDNA clones obtained from a 0–9-h embryo cDNA library (gift from K. Zinn) indicate that the open reading frame is contained within a single exon, that there is at least one 5' non-coding exon, and that the putative initiator methionine is the 5'-most proximal ATG codon and is located 5 codons 3' of an in-frame termination codon. The *Dfz2* and *Fz* sequences were aligned using GeneWorks software, and hydropathy profiles were calculated using MacVector 3.5 software. RNA from various stages of *Drosophila* development was isolated by the guanidinium–phenol method³⁸ or from tissue culture cells by LiCl–urea precipitation³⁹. RNA was transferred and hybridized as described³⁸.



sion is first clearly detected at stage 6, where it is found in all cells between roughly 15% and 70% of egg length, including the invaginating cells of the ventral furrow (Fig. 2a). An emerging stripe pattern is evident by early stage 8, and by stage 10 *Dfz2* expression is clearly seen in 15 stripes in the presumptive head and

trunk regions, in the posterior midgut primordium, in a subset of cells at the site of anterior midgut invagination, and in the procephalic lobe (Figs 2b, c). Beginning at stage 12, during germ-band shortening, *Dfz2* expression declines in the epidermis and increases in the midgut and visceral mesoderm (Fig. 2f). Beginning at stage 9 and continuing throughout embryogenesis, *Dfz2* expression is also seen in the developing CNS (Fig. 2d–l). By stage 17 *Dfz2* expression becomes limited to the CNS, hindgut and dorsal vessel (Fig. 2l).

Transfection of *Dfz2* in S2 cells

We also measured *Dfz2* RNA and protein levels in clone 8 cells, which respond to added Wg as determined by an increase in Arm protein, and in Schneider (S2) cells, which are unresponsive^{16,26}. A *Dfz2* transcript that matches the size of the *in vivo* transcript was observed in clone 8 cells, but no *Dfz2* transcript was detected in S2 cells (Fig. 1c). Similarly, in western blots probed with affinity-purified anti-*Dfz2* antibodies, a protein band with an apparent mobility of 65K (and comigrating with *Dfz2* produced in transfected S2 cells; see below) was observed in samples derived from clone 8 cells but not S2 cells (data not shown). Transcripts derived from *fz* were not detectable in either S2 or clone-8 cells (data not shown).

It is possible that S2 cells fail to respond to added Wg because they do not express *Dfz2*. We tested this directly by transfecting S2 cells with a *Dfz2* expression construct. These cells were then

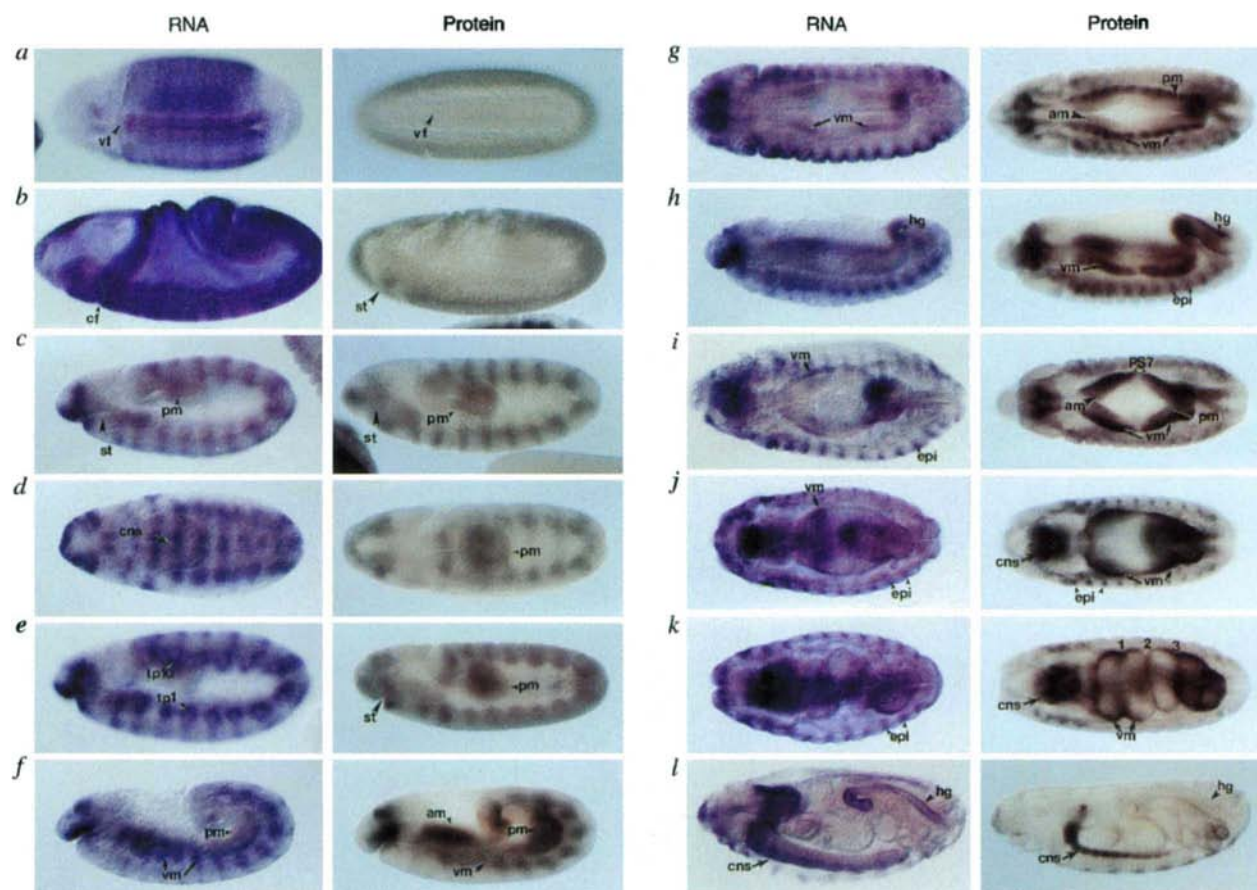


FIG. 2 *Dfz2* RNA and protein in wild-type embryos. Pairs of embryos at the same developmental stage are shown with *in situ* hybridization on the left (blue) and immunostaining on the right (brown). Embryos are oriented with anterior to the left. For embryos shown in a lateral view, dorsal is up. Stages and structures are according to ref. 40. *a*, Stage 6, dorsal/ventral view. *Dfz2* RNA is present in the central region of the embryo and is absent from the anterior and posterior regions. At this stage, *Dfz2* protein is below the limit of detection. *b*, Stage 7/8, lateral view. *Dfz2* RNA begins to accumulate in a striped pattern. *c*, Stage 9/10, lateral view. During germ-band extension, *Dfz2* is expressed in 15 stripes in the presumptive head and trunk regions, in the posterior midgut primordium, and in a subset of cells at the site of anterior midgut invagination. *d*, Stage 9/10, dorsal/ventral view. *Dfz2* expression can be seen in the developing CNS. *e*, Stage 10/11, lateral view. *Dfz2* expression resembles stage 9, with transiently higher expression around the primordia of the tracheal pits. *f*, Stage 12, lateral view. During germ-band retraction, *Dfz2* expression decreases in the epidermis but is maintained at high levels in the anterior and posterior midgut and the presumptive visceral mesoderm. *g*, Stage 13, dorsal/ventral view. The striped pattern of *Dfz2* expression persists in the visceral mesoderm and reappears in cells surrounding the segmental borders. *h*, Stage 13, lateral view. *Dfz2* is expressed at high levels in the hindgut. *i*, Stage 14, dorsal/ventral view. *Dfz2* expression is lower in the anterior and posterior midgut. There is ubiquitous expression in the visceral mesoderm, except in parasegment 7, previously described as the domain of *Ubx* and *dpp*

expression^{41,42}. *j*, Stage 15, dorsal/ventral view. *Dfz2* is expressed in the CNS, ventral mesoderm, and in cells surrounding the segmental borders. *k*, Stage 16, dorsal/ventral view. *Dfz2* is expressed at high levels in the CNS and in the visceral mesoderm spanning the first midgut constriction and posterior to the second midgut constriction. *l*, Stage 17, lateral view. *Dfz2* expression is primarily in the CNS, with lower levels in the hindgut and the dorsal vessel. Abbreviations: am, anterior midgut; cf, cephalic furrow; cns, central nervous system; epi, epidermis; hg, hindgut; pm, posterior midgut; PS 7, parasegment 7; st, stomodeum; tp, tracheal pit; vf, ventral furrow; vm, visceral mesoderm.

METHODS. Whole-mount embryo *in situ* hybridization was performed on 0–24-h embryos using a digoxigenin-labelled DNA probe encompassing *Dfz2* codons 1–307 as described⁴³ with minor variations. Fixation was in 4% formaldehyde/1× PBS and the staining reaction was done without levamisole. Identical patterns were obtained with a second probe corresponding to *Dfz2* codons 308–668. Immunohistochemical localization of *Dfz2* protein was done using affinity-purified rabbit antibodies raised against a fusion protein containing the bacteriophage T7 gene-10 protein joined to amino acids 65–314 of *Dfz2*. Antibodies were purified using a fusion protein containing the *E. coli* maltose-binding protein joined to the same segment of *Dfz2*. Immunostaining was done as described⁴⁴, except that embryos were fixed in Bouin's solution for 30 min instead of 4% formaldehyde/PBS. Antibody staining was visualized by the ABC method (Vector Labs) and embryos were mounted in methyl salicylate.

assayed for the ability to stabilize Arm in response to added Wg. In the absence of Wg, transfected and untransfected S2 cells show similar low levels of Arm irrespective of whether *Dfz2* expression was elevated by copper induction. However, when *Dfz2*-transfected cells were incubated in the presence of Wg, the level of the faster migrating (hypophosphorylated) form of Arm was increased (Fig. 3). This elevation was similar to the response elicited by Wg in clone-8 cells. Increasing *Dfz2* above basal level by copper induction of the metallothionein promoter led to a decrease in Wg responsiveness (Fig. 3), suggesting that at high levels *Dfz2* may bind non-productively to second messenger

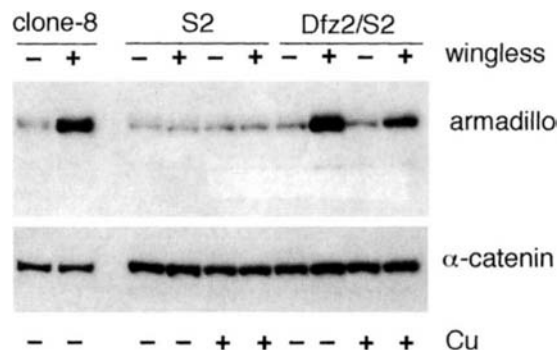
components. Four independent *Dfz2*-transfected cell lines derived from two separate transfections were tested, and all four lines showed Wg-dependent Arm stabilization.

Transfection of *Dfz2* confers Wg binding

The results described above show that *Dfz2* expression confers responsiveness to Wg, consistent with the idea that *Dfz2* is a receptor for Wg. To examine Wg binding directly, we incubated *Dfz2*-expressing S2 cells with Wg at 4 °C, and subsequently stained the cells with affinity-purified polyclonal antibodies to Wg. S2 cells expressing *Dfz2* show strong surface staining when incubated with

FIG. 3 Addition of soluble Wg leads to an increase in the level of Armadillo protein in *Dfz2*-transfected S2 cells. Clone 8 cells (left), untransfected S2 cells (centre), or *Dfz2*-transfected S2 cells (right) were incubated with concentrated conditioned medium either from S2 cells producing Wg (+Wg) or from control S2 cells (–Wg). Untransfected and *Dfz2*-transfected S2 cells were tested following growth with or without copper sulphate (+Cu or –Cu) to modulate expression of transfected *Dfz2* from the metallothionein promoter⁴⁵. Cellular proteins were analysed on blots with antibodies against Arm (upper panel). Incubation with Wg-containing medium produces an increase in the level of Arm in clone 8 cells and in *Dfz2*-expressing S2 cells, but not in untransfected S2 cells. Further induction of *Dfz2* expression in transfected cells by preincubation with copper sulphate leads to a lower response to Wg. As a control for loading, blots were stripped and incubated with antiserum against α -catenin (lower panel).

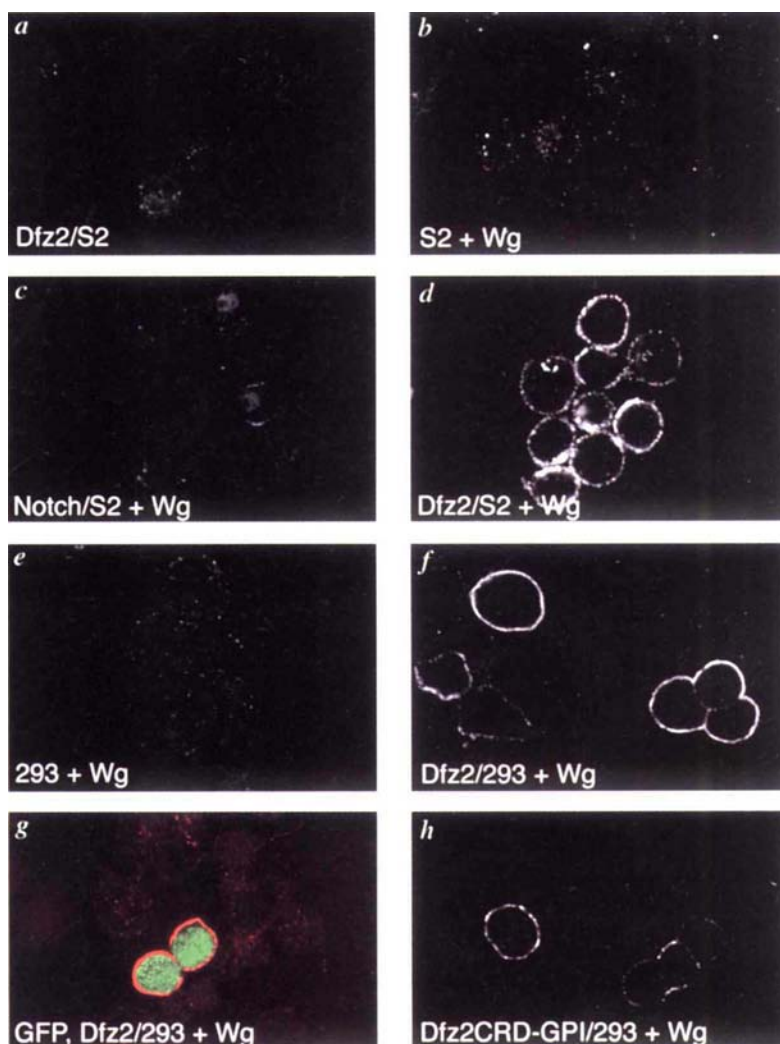
METHODS. To produce control conditioned medium or conditioned medium containing soluble Wg protein, untransfected S2 cells or S2 cells stably transfected with a construct in which the wg coding region is under the control of a heat-shock promoter were used as described⁴⁶. S2 cells stably transfected with *Dfz2* under the control of the metallothionein promoter were generated by hygromycin selection following transfection with a plasmid carrying the *Dfz2* coding region inserted into pMK33⁴⁵. Clone 8,



S2, and *Dfz2* transfected S2 cells were incubated with concentrated medium for 2 h. After the incubation, cells were lysed and protein extracts analysed using a monoclonal anti-armadillo antibody 7A1 (ref. 7) or rat-monoclonal anti- α -catenin antibody DCAT-1 (ref. 46). Bound antibody was visualized using the ECL system (Amersham).

FIG. 4 Wg protein binds to cells transfected with *Dfz2*. Untransfected and transfected cells were incubated with concentrated conditioned medium from untransfected S2 cells (a) or from S2 cells producing Wg (b–h; see Fig. 3 legend). Following incubation with conditioned medium, the cells were washed, fixed in paraformaldehyde/PBS, and incubated with an anti-Wg antibody directed against an 85-amino-acid domain that is found in Wg but absent from all other Wnt proteins⁴. This domain is dispensable for Wg activity (Chi-hwa Wu, C.H.S. and R.N., unpublished observations). Untransfected S2 cells (b), S2 cells transfected with a *Notch* expression plasmid (c), and untransfected 293T cells (human embryonic kidney-cell line 293 stably expressing SV40 TAG; e) show a low level of fluorescent antibody binding. d, Roughly 80% of S2 cells stably transfected with *Dfz2* and incubated with Wg show anti-Wg antibody binding to the cell surface. f, 10–20% of 293T cells transiently transfected with a *Dfz2* expression plasmid and incubated with Wg show anti-Wg antibody binding to the cell surface. g, 293T cells cotransfected with a mixture of *Dfz2* expression plasmid and a GFP expression plasmid, and incubated with Wg and anti-Wg antibodies show colocalization of green cytosolic fluorescence (GFP) and red surface fluorescence (Wg and anti-Wg antibody) and confirm that 10–20% of cells were transfected, and that only this subset of cells bind Wg. h, 293 cells cotransfected with a T-antigen expression plasmid and a truncated *Dfz2* construct in which the signal sequence, the CRD, and the first half of the linker region are anchored to the cell surface by GPI, show cell-surface binding to Wg and anti-Wg antibody.

METHODS. Untransfected S2 cells and S2 cells expressing *Dfz2* were washed twice in PBS and incubated with 1.5 ml of 10 $\times$ concentrated conditioned medium at 4 °C for 3 h. After three 10-min washes with cold PBS the cells were fixed in 2% paraformaldehyde (Polysciences, Inc.) for 15 min at room temperature. After three more 10-min washes with PBS, affinity purified anti-Wg antibody diluted 1:25 in 5% donkey serum/PBS was added to the cells and incubated overnight at 4 °C. After additional washes in PBS, the cells were incubated with fluorescent Cy3 secondary antibody (Jackson Immunoresearch) and mounted. For transient expression in 293T or 293 cells, the *Dfz2* coding region was inserted into the pCIS expression vector under the control of the cytomegalovirus immediate early promoter/enhancer and with an optimized translation-initiation context, and transfected into 293T or into 293 cells with a T-antigen expression plasmid using the calcium phosphate method⁴⁷. Eight hours after transfection, 10 mM chlorate was added. Twenty-four hours later, the cells were treated with 20 mU of heparitinase (Seikagaku) for 3 h before adding Wg protein. S2 cells expressing *Notch*⁴⁸ were obtained from S. Artavanis-Tsakonas. From N to C-termini, the GPI-anchored construct consists of the first 270 amino acids of *Dfz2*, a myc epitope⁴⁹ and the C-terminal 40 amino acids of decay activating factor, a GPI-anchored protein⁵⁰. Confocal images were collected with a Bio-Rad MRC 1000 confocal laser attached to a Zeiss Axio scope microscope. The same number of scans (20) were taken to visualize the fluorescence of each sample. Images were processed in Adobe Photoshop 3.0.



Wg and anti-Wg antibodies (Fig. 4d), whereas *Notch* transfected or non-transfected cells incubated either with or without Wg show a background of randomly distributed spots of low fluorescence intensity (Fig. 4b, c). *Dfz2*-transfected cells incubated in the absence of Wg show a similar low-intensity, spotty background (Fig. 4a). We conclude that Wg specifically binds to S2 cells expressing *Dfz2*.

Although this binding experiment indicates that Wg and *Dfz2* probably interact directly, it is possible that expression of *Dfz2* could act indirectly by inducing or unmasking a Wg receptor. We therefore performed a series of binding experiments using heterologous cells, in this case human embryonic kidney cells (293 or 293T; both will be referred to as 293) and a variety of wild-type and mutant fz constructs. Pretreatment of the 293 cells with chlorate and heparinase²⁷ lowered the overall background of Wg binding (presumably binding of Wg to extracellular matrix molecules; Fig. 4e), and revealed specific binding of Wg to the surface of 293 cells that had been transiently transfected with *Dfz2* (Fig. 4f) but not to untransfected cells or cells that had been transfected with a bovine rhodopsin expression construct (Fig. 4e, and data not shown). In a second experiment in which 293 cells were cotransfected with a green fluorescent protein (GFP) expression plasmid and the *Dfz2* expression plasmid, we observed that cells with green cytosolic fluorescence (caused by GFP) also had red surface fluorescence (Wg and anti-Wg antibody; Fig. 4g).

Each frizzled protein has an extracellular cysteine-rich domain (CRD) that is joined to the transmembrane domain by a variable linker. The CRD has been proposed to constitute part or all of the ligand-binding domain¹⁸, which suggests that cell-surface expression of the isolated CRD segment might confer Wg binding. This possibility was tested by expressing a truncated form of *Dfz2* in which the CRD and part of the linker region was displayed on the cell surface as a glycosylphosphatidylinositol (GPI)-anchored protein. This protein was detected at the surface of transfected cells by immunostaining either with antibodies directed against the *Dfz2* extracellular domain or with antibodies to a myc epitope tag that was engineered near the C terminus of the GPI-anchored protein (data not shown). When 293 cells transfected with the GPI-anchored *Dfz2* CRD were incubated with Wg and anti-Wg antibodies, strong surface staining was observed (Fig. 4h). We conclude from this experiment that the *Dfz2* CRD constitutes either all, or a significant part of, the ligand-binding domain.

Transfection of a subset of *frizzled* members

In many ligand–receptor systems a single ligand can bind to more than one species of receptor, or a single receptor can bind to more than one species of ligand, or both. Among G-protein-coupled receptors there are many examples of receptor subtypes that recognize the same ligand but differ in effector coupling, tissue distribution and pharmacology. In the fibroblast growth factor (FGF), platelet-derived growth factor (PDGF) and transforming growth factor (TGF)- β /activin/inhibin systems, tissue-culture experiments show that different receptors can bind to a single ligand and that different ligands can bind to a single receptor^{28–30}.

As a first step in examining the question of ligand–receptor specificity in the Wnt–frizzled system, we tested the ability of Wg to bind to 293 cells transfected with *Drosophila* fz and with six mammalian *frizzled* sequences¹⁸. 293 cells transfected either with fz, human fz5 (*Hfz5*), or mouse fz4, fz7 or fz8 (*Mfz4*, *Mfz7* and *Mfz8*) bind added Wg (Fig. 5a), whereas transfection with *Mfz3* and *Mfz6* did not confer Wg binding (Fig. 5b, c). As a complement to the *Dfz2* CRD GPI-anchor experiment, a derivative of *Mfz4* was constructed in which the CRD was replaced with a myc epitope. In transfected 293 cells, immunostaining with an anti-myc antibody, and western blotting with an antibody specific for the *Mfz4* C terminus show, respectively, that the CRD-deleted *Mfz4* protein accumulated at the cell surface and to the same percentage of membrane protein as full-length *Mfz4* (Fig. 5e and data not shown). However, CRD-deleted *Mfz4* did not confer Wg binding (Fig. 5d). This experiment further implicates the CRD as

an essential determinant of Wg binding. Figure 6 summarizes all of the Wg–frizzled binding experiments described above.

Discussion

The experiments reported here identify a second member of the *frizzled* family in *Drosophila*, *Dfz2*, and show by the following two criteria that it can function as a receptor for Wg. First, transfection of S2 cells with *Dfz2* confers Wg responsiveness as determined by an increase in cytoplasmic Arm concentration, and second, transfection with *Dfz2* confers cell-surface binding of Wg in both homologous (S2) and heterologous (293) cell systems. It is important to note that these data do not rule out the possibility that additional molecules in the conditioned medium might associate with Wg and participate in its binding to the receptor. They also do not rule out the possibility that *Dfz2* is part of a larger complex at the cell surface; in such a complex *Dfz2* would be necessary but may not be sufficient for binding and/or signal transduction. We note that these experiments do not support the proposal that *Notch* is the Wg receptor³¹, because *Notch*-

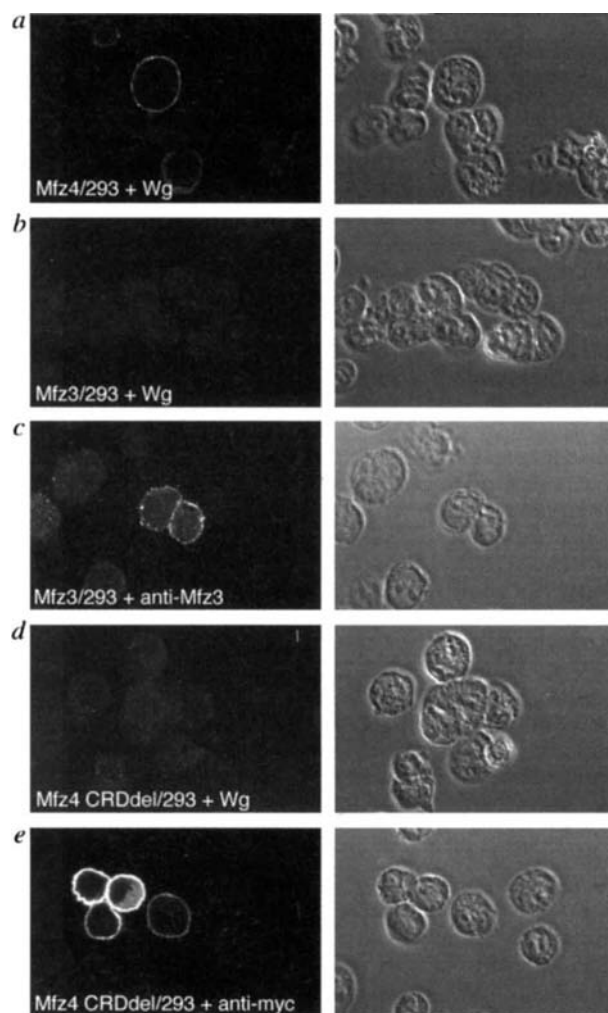


FIG. 5 Binding of Wg to 293 cells transfected with mammalian *frizzled* family members. 293 cells were cotransfected with a T-antigen expression plasmid and the following coding segments inserted into the pCIS vector: a, *Mfz4*; b, c, *Mfz3*; d, e, *Mfz4* with the CRD-replaced by a myc epitope. Cells were incubated with Wg and anti-Wg antibodies (a, b, d), affinity-purified antibodies directed against the extracellular domain of *Mfz3* (amino acids 1–205; c), or anti-myc antibodies (e). In each pair of photographs the left hand panel shows the immunostaining and the right hand panel the corresponding phase-contrast image.

METHODS. As described in Fig. 4.

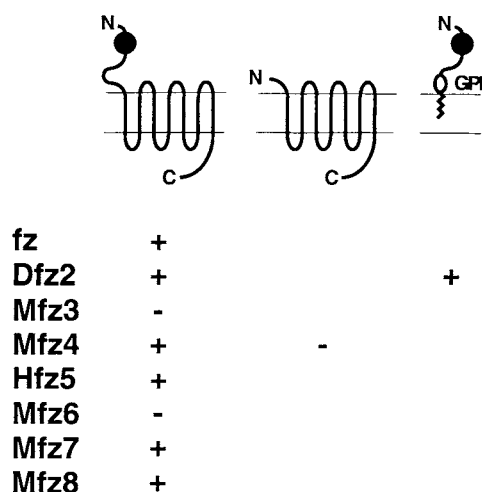


FIG. 6 Summary of Wg-frizzled interactions. Left, intact frizzled protein; centre, frizzled protein with the CRD deleted; right, the frizzled CRD and part of the linker region anchored to the membrane by GPI. + and - indicate the presence or absence of cell-surface binding by Wg after transfection of the frizzled proteins listed on the left. The filled ball represents the CRD. D, *Drosophila*; M, mouse; H, human. Whether Mfz6 is produced and transported to the cell surface has not been determined.

transfected cells do not bind Wg (Fig. 4c), nor does Notch confer a Wg-dependent increase in Arm (F. van Leeuwen and R.N., data not shown).

At present, there is no *in vivo* evidence that *Dfz2* is required for Wg signalling, as there are no known *Dfz2* mutants. Although the pattern of *Dfz2* expression is suggestive of its participation in wg signalling at multiple points in development, definitive evidence of that participation will require a genetic analysis of *Dfz2* function. The degree to which the *Dfz2* mutant phenotype resembles the wg phenotype will most probably depend on whether additional Wg receptors exist *in vivo*.

The ability of *Dfz2* to function as a Wg receptor implies more generally that other members of the *Wnt* and *frizzled* families are linked in receptor-ligand relationships. The observation that

Drosophila fz and some members of the mammalian *frizzled* family also confer Wg binding supports this inference but also suggests that there may be overlapping specificities in Wnt-Frizzled interactions. From the general conclusion that *frizzled* family members encode Wnt receptors, we infer that *in vivo* the *Drosophila fz* protein recognizes at least one Wnt other than Wg (three of which are known³²⁻³⁴), and, by extension, that the initial biochemical steps in fz-mediated tissue polarity signalling resemble the initial steps of wg-mediated segment polarity signalling.

The experiments reported here provide a new point of entry for examining the biochemistry of Wnt signalling. It should now be possible to determine which cytoplasmic proteins interact directly with the frizzled receptors, whether these interactions are modified by Wnt binding, and whether Wnt signalling is regulated by covalent or non-covalent receptor modification. It is interesting to note that many Fz proteins, including *Dfz2*, contain a S/T-X-V motif at their C-terminal end; this motif has been shown to interact with PDZ (or DHR) domains in a variety of proteins³⁵. Dsh, one of the cytoplasmic components of Wg signalling, contains a PDZ domain^{9,10}.

A question remaining is how frizzled and Wnt proteins might interact to initiate signal transduction. One attractive hypothesis is suggested by the relative immobility of Wnt proteins because of their affinity for the extracellular matrix, and the predicted mobility of the CRD, which we show here constitutes part or all of the ligand-binding site. The prediction that the CRD is mobile follows from the predicted lack of a stable structure in the highly divergent sequence that links it to the membrane-embedded domain. For example, in *Dfz2* this linker region includes a stretch of 42 amino acids that includes 21 glycines and 15 serines. Therefore the CRD may be able to bind to an extracellular matrix-associated Wnt protein at a distance of several tens of nanometres from the plasma membrane of the cell on which the frizzled receptor resides. It is tempting to speculate that binding of a Wnt ligand to the CRD disrupts or modifies an interaction between the CRD and the extracellular face of the transmembrane domain, and that this results in a rearrangement of transmembrane α -helices. Although the frizzled proteins have no primary sequence homology to G-protein-coupled receptors¹⁸, this allosteric model suggests a mode of receptor activation that is reminiscent of that proposed for G-protein-coupled receptors³⁶. □

Received 29 April; accepted 24 June 1996.

- Nusse, R. & Varmus, H. E. *Cell* **69**, 1073-1087 (1992).
- Klingensmith, J. & Nusse, R. *Dev. Biol.* **166**, 396-414 (1994).
- Herman, M. A. & Horvitz, H. R. *Development* **120**, 1035-1047 (1994).
- Rijsewijk, F. et al. *Cell* **50**, 649-657 (1987).
- Baker, N. E. *EMBO J.* **6**, 1765-1773 (1987).
- Perrimon, N. *Cell* **76**, 781-784 (1994).
- Peifer, M. *J. Cell Sci.* **105**, 993-1000 (1993).
- Bejsovec, A. & Wieschaus, E. *Development* **119**, 501-517 (1993).
- Klingensmith, J., Nusse, R. & Perrimon, N. *Genes Dev.* **8**, 118-130 (1994).
- Theisen, H. et al. *Development* **120**, 347-360 (1994).
- McCrea, P. D., Turck, C. W. & Gumbiner, B. *Science* **254**, 1359-1361 (1991).
- Siegrfried, E., Wilder, E. L. & Perrimon, N. *Nature* **367**, 76-80 (1994).
- Peifer, M., Sweeton, D., Casey, M. & Wieschaus, E. *Development* **120**, 369-380 (1994).
- Noordermeer, J., Klingensmith, J., Perrimon, N. & Nusse, R. *Nature* **367**, 80-83 (1994).
- Riggelman, B., Schedl, P. & Wieschaus, E. *Cell* **63**, 549-560 (1990).
- Van Leeuwen, F., Harryman Samos, C. & Nusse, R. *Nature* **368**, 342-344 (1994).
- Peel, D. J. & Milner, M. J. *Wilhelm Roux's Arch. dev. Biol.* **201**, 120-123 (1992).
- Wang, Y. et al. *J. Biol. Chem.* **271**, 4468-4476 (1996).
- Adler, P. N., Charlton, J. & Vinson, C. *Dev. Biol.* **8**, 99-119 (1987).
- Gubb, D. & Garcia, B. A. *J. Embryol. exp. Morphol.* **68**, 37-57 (1982).
- Zheng, L., Zhang, J. J. & Carthew, R. W. *Development* **121**, 3045-3055 (1995).
- Vinson, C. R., Conover, S. & Adler, P. N. *Nature* **338**, 263-264 (1989).
- Adler, P. N., Vinson, C., Park, W. J., Conover, S. & Klein, L. *Genetics* **126**, 401-416 (1990).
- Vinson, C. R. & Adler, P. N. *Nature* **329**, 549-551 (1987).
- Krasnow, R. E., Wong, L. L. & Adler, P. N. *Development* **121**, 4095-4102 (1995).
- Yanagawa, S., van Leeuwen, F., Wodarz, A., Klingensmith, J. & Nusse, R. *Genes Dev.* **9**, 1087-1097 (1995).
- Olwin, B. B. & Rapraeger, A. J. *Cell Biol.* **118**, 631-639 (1993).
- Johnson, D. E. & Williams, L. T. *Adv. Cancer Res.* **60**, 1-41 (1993).
- Seifert, R. A. et al. *J. Biol. Chem.* **264**, 8771-8778 (1989).
- Ten Dijke, P., Miyazono, K. & Heldin, C. H. *Curr. Opin. Cell Biol.* **8**, 139-145 (1996).
- Couso, J. P. & Martinez Arias, A. *Cell* **79**, 259-272 (1994).

- Russell, J., Gennissen, A. & Nusse, R. *Development* **115**, 475-485 (1992).
- Graba, Y. et al. *Development* **121**, 209-218 (1995).
- Eisenberg, L. M., Ingham, P. W. & Brown, A. M. C. *Dev. Biol.* **154**, 73-83 (1992).
- Gomperts, S. N. *Cell* **84**, 659-662 (1996).
- Khorana, H. J. *biol. Chem.* **267**, 1-4 (1992).
- Kyte, J. & Doolittle, R. J. *molec. Biol.* **157**, 105-132 (1982).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, NY, 1989).
- Auffray, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303-314 (1980).
- Campos-Ortega, J. A. & Hartenstein, V. *The Embryonic Development of Drosophila melanogaster* (Springer, Berlin, 1985).
- Mathies, L. D., Kermidge, S. & Scott, M. P. *Development* **120**, 2799-2809 (1994).
- Bienz, M. *Trends Genet.* **10**, 22-26 (1994).
- Tautz, D. & Pfeifle, C. *Chromosoma* **98**, 81-85 (1989).
- Reuter, R., Pangniban, G. E. F., Hoffmann, F. M. & Scott, M. P. *Development* **110**, 1031-1040 (1990).
- Koelle, M. R. et al. *Cell* **67**, 59-77 (1991).
- Oda, H. et al. *J. Cell Biol.* **121**, 1133-1140 (1993).
- Gorman, C., Gies, D. & McCray, G. *DNA Prot. Engng Technol.* **2**, 3-9 (1990).
- Fehon, R. G. et al. *Cell* **61**, 523-534 (1990).
- Evan, G. I., Lewis, G. K., Ramsay, G. & Bishop, J. M. *Molec. cell. Biol.* **5**, 3610-3616 (1985).
- Caras, I. W. & Weddell, G. N. *Science* **243**, 1196-1198 (1989).

ACKNOWLEDGEMENTS. P.B. and M.B. contributed equally to this work. We thank P. Adler, S. Artavanis-Tsakonas, I. Caras, C. Machado, A. Rattner, H. Thaker, A. Spradling, C. Thummel and K. Zinn for advice and materials; A. Wodarz and M. Delannoy for help with using the confocal microscopes; and P. Beachy, D. Kingsley, J. Nelson, G. Barsh, M. Scott and members of the Nathans and Nusse laboratories for discussions and comments on the manuscript. M.B. was supported by a long-term fellowship from EMBO. These studies were supported by the Howard Hughes Medical Institute, in which J.N. and R.N. are investigators, and by a grant to R.N. from the USAMRMC.

CORRESPONDENCE and requests for materials should be addressed to J.N. or R.N. (e-mail: jeremy.nathans@qmail.bs.jhu.edu or nusse@cmm.stanford.edu).

Detection of Lyman- α -emitting galaxies at redshift 4.55

Esther M. Hu* & Richard G. McMahon†

*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

†Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

STUDIES of the formation and early history of galaxies have been hampered by the difficulties inherent in detecting faint galaxy populations at high redshift. As a consequence, observations at the highest redshifts ($z = 3.5$ – 5) have been restricted to objects that are intrinsically bright. These include quasars, radio galaxies, and some Lyman- α -emitting objects^{1–3} that are very close to (within ~ 10 kpc)—and appear to be physically associated with—quasars. But the extremely energetic processes which make these objects easy to detect also make them unrepresentative of normal (field) galaxies. Here we report the discovery of two Lyman- α -emitting galaxies at redshift $z = 4.55$, which are sufficiently far from the nearest quasar (~ 700 kpc) that radiation from the quasar is unlikely to provide the excitation source of the Lyman- α emission. Instead, these galaxies appear to be undergoing their first burst of star formation, at a time when the Universe was less than one billion years old.

Searches for high-redshift galaxies have progressed from several directions. Extensive spectroscopic surveys of field galaxies^{4–7} now show numerous random field galaxies at $z > 1$, with evidence that many of these are in the process of intense star formation⁸. Searches for the galaxies associated with quasar absorption lines are also beginning to find normal galaxies in this redshift range⁹. At $z > 1.7$, where the [O II] emission line at 3,727 Å is no longer detectable in the optical window, we are beginning to find objects on the basis of their ultraviolet absorption lines^{10–13}. These objects may be starbursting galaxies, in some cases with active galactic nuclei (AGN).

At redshifts near 3, absorption-line objects should also show distinct continuum colour breaks as the limit of the hydrogen Lyman series (912 Å in the rest frame) becomes observable at ultraviolet wavelengths. The objects identified by colour breaks and ultraviolet absorption features to lie at $3.0 < z < 3.5$ (ref. 13) should be relatively evolved galaxies, whose bright ultraviolet continua are produced by massive stars which chemically enrich these systems, giving rise to metal-line absorption features and dust, making Ly α emission weak or undetectable.

Galaxies at earlier stages of this process, before the formation of dust, may have much stronger Ly α emission relative to the stellar continuum. Such objects, which are faint in the continuum, may be hard to pick out with colour-break techniques but relatively easy to find by Ly α searches. A small number of such objects have been seen in the fields of quasars with redshifts 2–3.5, which are well separated from the quasar and which may represent neighbouring galaxies either with AGN¹⁴ or with intense star

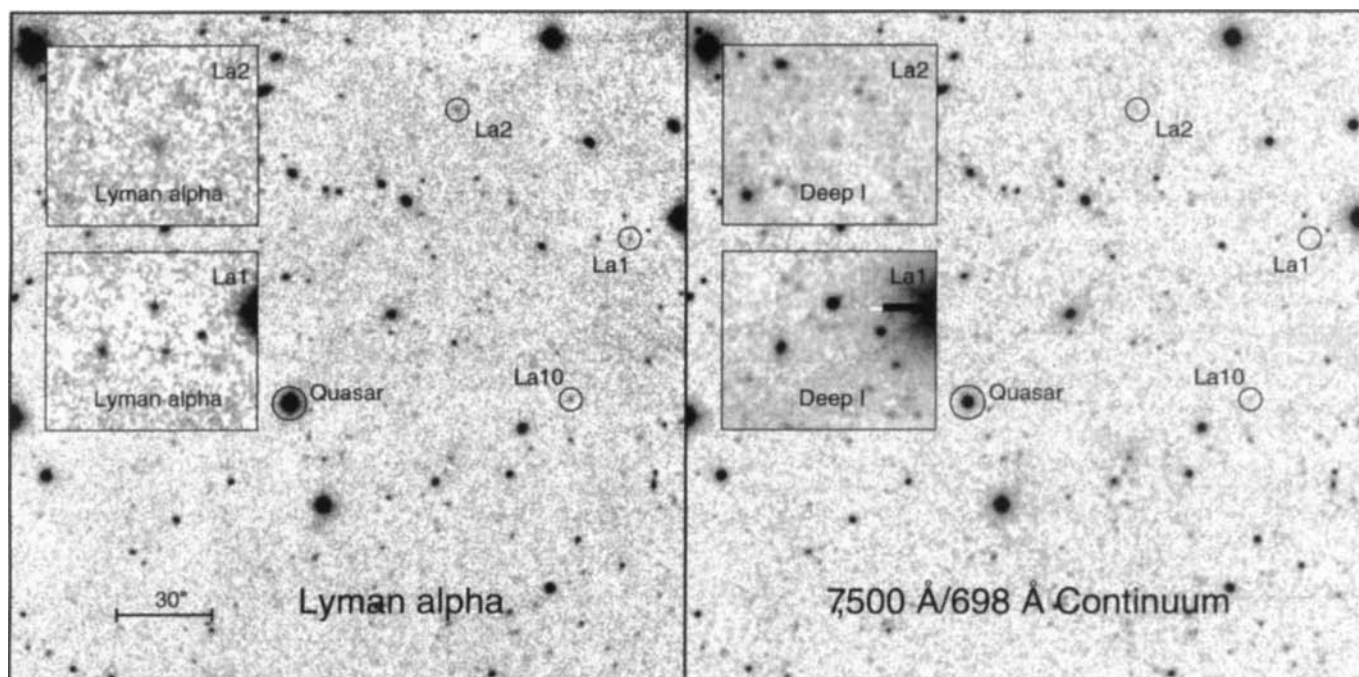


FIG. 1 Central 3.5×3.5 arcmin region showing the $z = 4.55$ quasar BR2237 – 0607 and emitting objects LA1, LA2 and LA10 in a narrow-band filter (central wavelength 6,741 Å, 51-Å bandpass) centred on the quasar's redshifted Ly α emission (left panel) and through a nearby broad, line-free continuum filter (7,500 Å, 698-Å bandpass) (right panel). Inset panels show the region around identified Ly α emitters LA1 and LA2 both in line and continuum, and enlarged by a factor of two. Here the continuum is a deep one-hour I-band taken at Keck. LA2 has a weak continuum (Fig. 2), but is clearly extended in both the Ly α and I-band images. The continuum for LA1 is not detected in the one-hour I-band exposure at a 1σ magnitude of 26.15. The data were taken as a series of sky-noise-limited integrations,

with an offset step of 10 arcsec between successive frames, and a median sky exposure was generated from the on-field exposures for each night and used to correct for non-uniformities in detector response (flat-field). Spectrophotometric standards (Feige 15, Feige 110, Kopff 27, BD + 28 4211)^{28,29} for the narrow-band data and a combination of spectrophotometric (Feige 110, BD + 28 4211)²⁹ and Landolt standards (in the fields of SA 92-248, PG1633 + 099, PG0231 + 051)³⁰ for the continuum data were used to calibrate the data, and were taken in observations both before and after the target exposures. The composite image quality was 1.0 arcsec in both the 20.8-h narrow-band image and in the 1-h 7,500 Å continuum image, and 0.8 arcsec in the deep 1-h Keck I-band image.

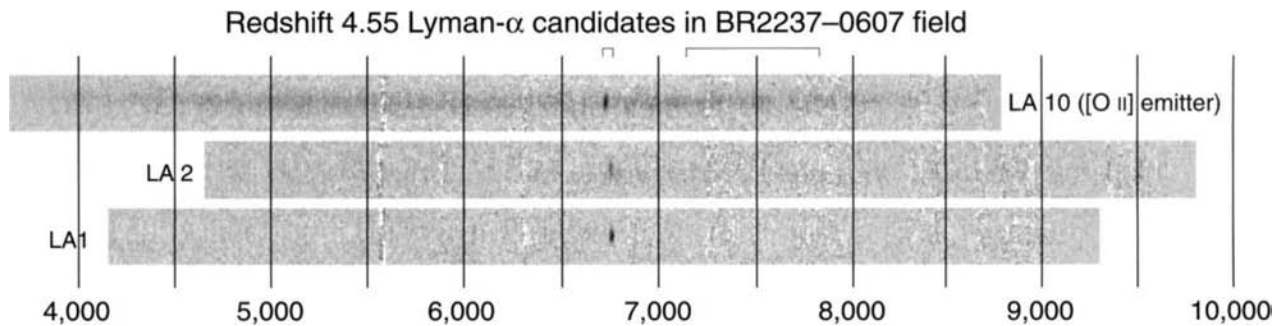


FIG. 2 LRIS spectra of the three candidates from Fig. 1 in the BR2237 – 0607 field. Observations were made with a multi-slit mask over the quasar field, with a slit width of 1.4 arcsec, spectral resolution of ~ 17 Å, and spatial sampling of 0.215 arcsec per pixel. The results of 3 h exposure in a single mask setting are shown as two-dimensional spectra, with night sky lines subtracted from the spectrum, and the displaced starting wavelengths reflect the offsets in slit position for different objects in the field. The spatial extent of each spectrum is ~ 6.5 arcsec—roughly the diameter encircling the objects in Fig. 1. At the top of the plot we indicate the wavelength regions covered by the narrow-band filter (central wavelength 6,741 Å, 51 Å bandpass), which is centred near the wavelength of the

quasar's Ly α emission, and the reference continuum filter (central wavelength 7,500 Å, 698 Å bandpass). Objects LA1 and LA2, which are well centred in the narrow-band filter, have observed equivalent widths in excess of 700 Å in the emission line, and thus must almost certainly be due to redshifted Ly α . No other features are detected, apart from a faint continuum in LA2. Object LA10 is an [O II] emitter at $z = 0.802$ with a line equivalent width of ~ 123 Å, and can be easily identified, even in the absence of other emission lines, by its strong continuum, extending to the blue of 5,000 Å (roughly, the wavelength of the break in the Lyman-limit absorption for $z \approx 4.6$ objects). The asymmetric placement of the emission line is another sign of a low-redshift interloper (see text).

formation that can excite the Ly α emission line but without so much dust that the line is suppressed^{15–17}. The latter class of objects, which are relatively unobscured by foreground absorption, may provide the most direct estimates of star formation at the earliest epochs. It appears that the quasar may mark the sites of such formation and thus identify a redshift for targeted wavelength searches¹⁸. This Letter reports detection of such objects at by far the highest redshift yet in the field of the $z = 4.55$ quasar BR2237 – 0607¹⁹.

To search for high-redshift objects in the fields of the $z > 4$ quasars, we first obtained exposures through a narrow-band filter centred on the quasar's redshifted Ly α emission, and through a nearby broader continuum, with a Tektronix 2048² camera on the University of Hawaii's 2.2-m telescope. Five $z > 4$ quasar fields have now been imaged, but the spectroscopic follow-up is complete only for BR2237 – 0607, which we describe here. Narrow-band and continuum images on this field are shown in Fig. 1. All objects with a significant excess in the Ly α bandpass were considered candidate Ly α emitters, and this initial search found 10 objects in the 4.39×4.39 arcmin field surrounding BR2237 – 0607, a subsample of which is circled in Fig. 1. However, many of these candidate objects correspond to lower-redshift galaxies whose emission lines coincidentally match the redshifted Ly α wavelength, or to objects whose continuum

shape makes them significantly brighter in the narrow-band filter. Therefore, to proceed further we obtained simultaneous spectra of the candidate objects using the multi-object LRIS spectrograph²⁰ on the Keck 10-m telescope.

Two-dimensional spectra for the three candidate objects of Fig. 1 are shown in Fig. 2, with the wavelength region of the narrow-band and reference continuum-band filters marked at the top of the plot. As expected, most of the candidate objects were identified as $z < 1$ galaxies. We show in this figure, for instance, the spectrum of the low-redshift emission-line galaxy, LA10, where the strong emission is identified with the [O II] 3,727 Å line instead of Ly α , for a redshift of 0.802. The spectra show various features that are hallmarks of this class of interloper, namely: much stronger continuum, including clearly detected flux from the region below 5,000 Å which would correspond to the Lyman-limit break at the rest-frame 912 Å photoionization threshold of neutral hydrogen for $z \approx 4.6$ objects; and the offset wavelength position near the edge of the narrow passband filter. However, the two objects denoted here as LA1 and LA2 exhibit only an extremely strong single emission line close to the central wavelength of the narrow-band filter together with an extremely weak continuum. The insets in Fig. 1 show that for LA1 the continuum is undetected even in an extremely deep one-hour I-band image obtained with LRIS on the Keck 10-m telescope, while a weak

TABLE 1 BR2237 – 0607 candidate identifications

Object	z	RA (1950)	Dec. (1950)	Flux (6,741 Å) ($\text{erg cm}^{-2} \text{s}^{-1}$)	m_{6741}	m_{7500}	m_i
Quasar	4.558*	22 h 37 min 17.40 s	–06° 07' 59.7"	2.4×10^{-14}	16.57	17.99	...
LA1	4.551*	22 h 37 min 10.35 s	–06° 07' 08.6"	5.0×10^{-17}	23.60	>24.89	>26.15
LA2	4.550*	22 h 37 min 13.92 s	–06° 07' 28.2"	6.7×10^{-17}	23.06	24.81	24.69
			(1σ)	1.5×10^{-17}	24.92	24.89	26.15

Object coordinates were determined by applying astrometric fits for plate scale and rotation, and fitting to the Cambridge Automatic Plate Measuring (APM) Survey object positions in the quasar field and in calibration star fields taken before and after the quasar observations. We estimate the relative position accuracy of these coordinates to have typical errors less than 0.2–0.3 arcsec. Magnitudes, fluxes and their 1σ errors are measured over 3 arcsec diameter apertures. The narrow-band and adjacent continuum-band filter measurements (m_{6741} and m_{7500}) are given as equivalent visual (AB) magnitudes, for ease of conversion to emission-line fluxes, which are quoted with continuum contributions subtracted. Keck I-band magnitudes are on the Kron-Cousins system, and were tied to I-band photometry on this field using the University of Hawaii 2.2-m telescope (Keck I zero point = 27.46). The quasar is saturated on the deep Keck I-band image.

* From Storrie-Lombardi *et al.*'s estimate (ref. 19) based on O I 1,304 Å, Si/O IV 1,400 Å and C IV 1,549 Å. Simultaneous observations of the quasar were not possible without dropping one or more of the potential candidates from the mask.

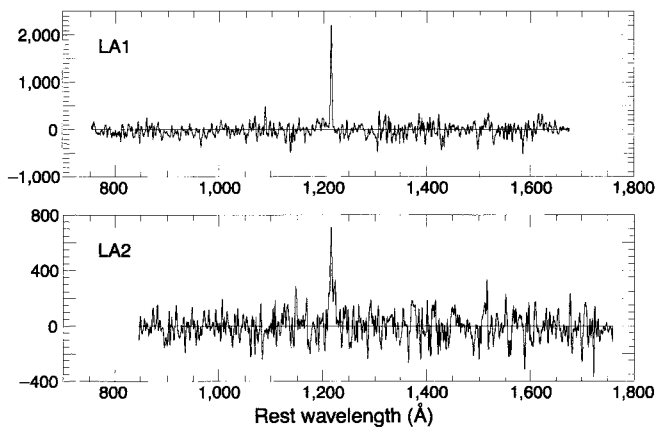


FIG. 3 Extracted spectra for LA1 and LA2 plotted in the rest-wavelength frame. The spectra represent a total of 4.7 h integration with the LRIS spectrograph on the Keck telescope. The emission wavelengths agree to within $\sim 2.4 \text{ \AA}$ ($\sim 100 \text{ km s}^{-1}$ at $z = 4.55$; $\sim 0.4 \text{ \AA}$ in the rest frame) for these two objects. LA1 is not resolved at $3.1 \text{ \AA}$ in the rest frame, and the width of LA2 is probably due to the extended emission structure across the slit.

diffuse structure is seen in this band for LA2, whose continuum is also faintly detected in the spectroscopic data. The Kron–Cousins I magnitude of LA2 is 24.69 and it is detected at the 4σ level in the Keck I -band image (Table 1). Spectra of these two objects are shown in Fig. 3. A useful measure of the relative strength of line to continuum is the equivalent width, which is defined as the width of a rectangle with the same area as an emission (or absorption) feature, with a height equal to the continuum level above zero. The observed equivalent widths ($>700 \text{ \AA}$ (LA2) and $>1,300 \text{ \AA}$ (LA1)) are quoted as lower limits because of the weakness of the continuum, but are too large for these lines to be $[\text{O II}]$ or $\text{H}\alpha$ (refs 21–23)—an interpretation which is also ruled out by the absence of any other strong emission lines. This leaves little alternative but to identify these objects as $\text{Ly}\alpha$ emitters at the redshift of the quasar. Their coordinates and properties may be found in Table 1. The equivalent width of object LA10 ($123 \text{ \AA}$) is consistent with the expected range of values for $[\text{O II}]$ emission^{4,7}.

The separations from the quasar of 117 arcsec for LA1 and 105 arcsec for LA2 are too large for the quasar to have a significant role in exciting the objects²⁴ ($1 \text{ arcsec} = 7.1 h^{-1} \text{ kpc}$ at $z = 4.55$, where $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.5$) and it appears that the objects must be either internally excited or ionized by the general metagalactic ultraviolet flux²⁵. As can be seen from Fig. 1 inset, LA1 is quite compact and could have an AGN component, although there is no sign of emission from C IV at $1,549 \text{ \AA}$. However, LA2 is quite diffuse in both I and $\text{Ly}\alpha$, and probably excited by star formation. An alternative possibility is that the $\text{Ly}\alpha$ emission is produced by surface ionization of the gas cloud by the metagalactic ultraviolet flux. The metagalactic ionizing flux at these wavelengths is currently estimated²⁵ as $(1\text{--}3) \times$

$10^{-22} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}$, which would produce an observed $\text{Ly}\alpha$ flux of $\sim 2 \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$ from an object with a 2 arcsec angular size at $z = 4.55$ (ref. 23) if there is one $\text{Ly}\alpha$ photon emitted per ionization and we see only the forward side of the cloud, so it is possible we are seeing a component due to ionized gas clouds. However, this would provide no explanation for the continuum light in LA2.

The rest-frame equivalent widths of the systems are around $>130 \text{ \AA}$ (LA2) and $>240 \text{ \AA}$ (LA1), which are marginally consistent with stellar excitation for an initial mass function dominated by massive stars²⁶. Following the discussion of the previous paragraph, the $\text{Ly}\alpha$ may arise from a combination of internal and external ionization. If the observed emission were primarily due to stars, and there were no internal scattering and extinction, then the luminosity of $3 \times 10^{42} h^{-2} \text{ erg s}^{-1}$ ($q_0 = 0.5$) would correspond to a star formation rate in solar masses ($M_\odot$) per year of $\sim 3 h^{-2} M_\odot \text{ yr}^{-1}$, where we use Kennicutt's (ref. 21) relation between $\text{H}\alpha$ luminosity and star formation rate (SFR) of $\text{SFR} = L(\text{H}\alpha) \times 8.9 \times 10^{-42} \text{ erg s}^{-1} M_\odot \text{ yr}^{-1}$, and assume a ratio of $\text{Ly}\alpha$ to $\text{H}\alpha$ (8.7) that applies for case B recombination²⁷. Given the Hubble time at this redshift of $7 \times 10^8 h^{-1} \text{ yr}$ ($q_0 = 0.5$) the integrated amount of star formation is small compared to that of a 'normal' galaxy with $6 \times 10^{10} h^{-1} M_\odot$ of stars (a so-called L^* galaxy).

Because of the targeted nature of the search it is hard to estimate from the present data whether such objects may be common in the general field or whether they are preferentially found around quasars. Observations, currently in progress, of additional quasars and blank field regions should answer this question. $\square$

Received 4 March; accepted 4 June 1996.

- Hu, E. M., McMahon, R. G. & Egami, E. *Astrophys. J.* **459**, L53–L57 (1996).
- Fontana, A., Cristiani, S., D'Odorico, S., Giallongo, E. & Savaglio, S. *Mon. Not. R. astr. Soc.* **279**, L27–L30 (1996).
- Petitjean, P., Pécontal, E., Valls-Gabaud, D. & Charlot, S. *Nature* **380**, 411–413 (1996).
- Songaila, A., Cowie, L. L., Hu, E. M. & Gardner, J. P. *Astrophys. J. Suppl. Ser.* **94**, 461–515 (1994).
- Lilly, S. J., LeFèvre, O., Crampton, D., Hammer, F. & Tresse, L. *Astrophys. J.* **455**, 50–59 (1995).
- Glazebrook, K. et al. *Mon. Not. R. astr. Soc.* **273**, 157–168 (1995).
- Cowie, L. L., Songaila, A., Hu, E. M. & Cohen, J. G. *Astr. J.* (in the press).
- Cowie, L. L., Hu, E. M. & Songaila, A. *Nature* **377**, 603–605 (1995).
- Dickinson, M. & Steidel, C. *Bull. Am. astr. Soc.* **27**, 849 (1995).
- Cowie, L. L. et al. *Astrophys. J.* **432**, L83–L86 (1994).
- Egami, E., Iwamura, F., Maihara, T., Oya, S. & Cowie, L. L. *Astr. J.* (in the press).
- Yee, H. K. C., Ellingson, E., Bechtold, J., Carlberg, R. G. & Cuillandre, J.-C. *Astr. J.* (in the press).
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M. & Adelberger, K. L. *Astrophys. J.* **462**, L17–L21 (1996).
- Lowenthal, J. D. et al. *Astrophys. J.* **377**, L73–L77 (1991).
- Giavalisco, M., Macchetto, F. D. & Sparks, W. B. *Astr. Astrophys.* **288**, 103–121 (1994).
- Macchetto, F. D., Lipari, S., Giavalisco, M., Turshek, D. A. & Sparks, W. B. *Astrophys. J.* **404**, 511–520 (1993).
- Francis, P. J. et al. *Astrophys. J.* **457**, 490–499 (1996).

- Djorgovski, S., Spinrad, H., McCarthy, P. & Strauss, M. *Astrophys. J.* **299**, L1–L5 (1985).
- Storie-Lombardi, L. J., McMahon, R. G., Irwin, M. J. & Hazard, C. *Astrophys. J. Suppl. Ser.* (in the press).
- Oke, J. B. et al. *Publ. astr. Soc. Pacif.* **107**, 375–385 (1995).
- Kennicutt, R. C. Jr *Astrophys. J.* **272**, 54–67 (1983).
- Kennicutt, R. C. Jr *Astrophys. J.* **388**, 310–327 (1992).
- Songaila, A., Cowie, L. L. & Lilly, S. J. *Astrophys. J.* **348**, 371–377 (1990).
- Hu, E. M., Songaila, A., Cowie, L. L. & Stockton, A. *Astrophys. J.* **368**, 28–39 (1991).
- Williger, G. M. et al. *Astrophys. J.* **428**, 574–590 (1994).
- Charlot, S. & Fall, S. M. *Astrophys. J.* **415**, 580–588 (1993).
- Brocklehurst, M. *Mon. Not. R. astr. Soc.* **153**, 471–490 (1971).
- Massey, P., Strobel, K., Barnes, J. V. & Anderson, E. *Astrophys. J.* **328**, 315–333 (1988).
- Stone, R. P. S. *Astrophys. J.* **218**, 767–769 (1977).
- Landolt, A. U. *Astr. J.* **104**, 340–371 (1992).

ACKNOWLEDGEMENTS. E.M.H. is a Visiting Astronomer, W. M. Keck Observatory. We thank T. Bida, R. Campbell, T. Chelminiak, and B. Schaefer for their assistance in obtaining the observations, which would not have been possible without the LRIS spectrograph of J. Cohen and B. Oke. Research at the University of Hawaii was supported by the State of Hawaii and by NASA. E.M.H. acknowledges a University Research Council Seed Money grant. R.G.M. acknowledges the support of the Royal Society.

CORRESPONDENCE should be addressed to E.M.H. (e-mail: hu@ifa.hawaii.edu).

Identification of a galaxy responsible for a high-redshift Lyman- α absorption system

S. G. Djorgovski*, M. A. Pahre*, J. Bechtold† & R. Elston‡

* Palomar Observatory, MS 105–24, California Institute of Technology, Pasadena, California 91125, USA

† Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

‡ Cerro Tololo Interamerican Observatory, NOAO, Casilla 603, La Serena, Chile

DAMPED Lyman- α systems are high-column-density intergalactic clouds of hydrogen, the existence of which is inferred from absorption lines appearing in the emission spectra of distant quasars. The galaxies believed to be responsible for these absorption systems have been suggested as possible progenitors of the normal disk galaxies observed in the local Universe¹. Indeed, Lyman- α systems appear to contain a substantial fraction of the baryons known to exist in galaxies today^{2,3}. Here we report the optical detection of a galaxy (designated DLA2233 + 131) associated with a known damped Lyman- α absorption system at a redshift of $z = 3.150$. The properties of this galaxy correspond closely to those expected of a young disk galaxy in the early stages of formation, and show no evidence for an active nucleus. This finding gives strong support to the idea that damped Lyman- α systems represent a population of young galaxies at high redshifts.

Despite considerable efforts, no direct counterpart of a damped Lyman- α (DLA) galaxy has been unambiguously detected so far^{5–8}. Clustered companions of DLA objects, all containing active galactic nuclei (AGN), or quasar companions which may be responsible for some associated absorption have been seen, but no normal, isolated DLA systems themselves^{9–13}. The closest may be the objects apparently associated with a quasar PKS0528 – 250, with absorption redshift $z_{\text{abs}} = 2.811 > z_{\text{QSO,em}} = 2.77$ (where $z_{\text{QSO,em}}$ is the emission redshift of the quasar) but given their association with the quasar, their nature and the source of ionization remain uncertain^{12,13}.

We searched for the optical counterpart to DLA2233 + 131 in the field of the quasar 2233 + 131 ($z = 3.295$, ref. 14). Our data were obtained on the night of 28 September 1995 UT using the Low Resolution Imaging Spectrograph (LRIS) instrument¹⁵ at the W. M. Keck 10-m telescope on Mauna Kea, Hawaii. The object was found during observations of a candidate for a different DLA system (at $z_{\text{abs}} = 2.551$)⁴ in the same field, which will be reported elsewhere. It was also selected independently as a DLA candidate on the basis of its broad-band colours by Steidel *et al.*¹⁶, which was unknown to us at the time. We obtained direct images of the field in the Cousins V and R bands, and long-slit spectra using low- and moderate-resolution gratings, providing spectroscopic resolutions of full-width at half-maximum (FWHM) ~ 10 and $3.5 \text{ \AA}$, respectively. The seeing was FWHM ~ 0.7 arcsec. Because no photometric standards were observed that night, V and R images of the same field were obtained in photometric conditions at the Palomar 60-inch telescope on the night of 17 October 1995 UT and used to establish the photometric zero point for the Keck images. All data were reduced using standard techniques.

The object is identified with a galaxy 2.3 arcsec away in projection from the quasar, in position angle 159° (Fig. 1). Its magnitudes are $V = 25.1 \pm 0.2$ and $R = 24.8 \pm 0.2$, in excellent agreement with the photometry by Steidel *et al.*¹⁶. The optical spectrum to the blue of the Lyman break is thus nearly flat, with flux $\langle F_\nu \rangle \approx 0.35 \mu\text{Jy}$, typical of actively star forming galaxies. For the quasar itself we obtain $V = 18.29$ mag and $R = 18.15$ mag.

Our spectra show a prominent Ly α line emission at redshift $z_{\text{em}} = 3.1530 \pm 0.0003$, corresponding to the rest-frame velocity

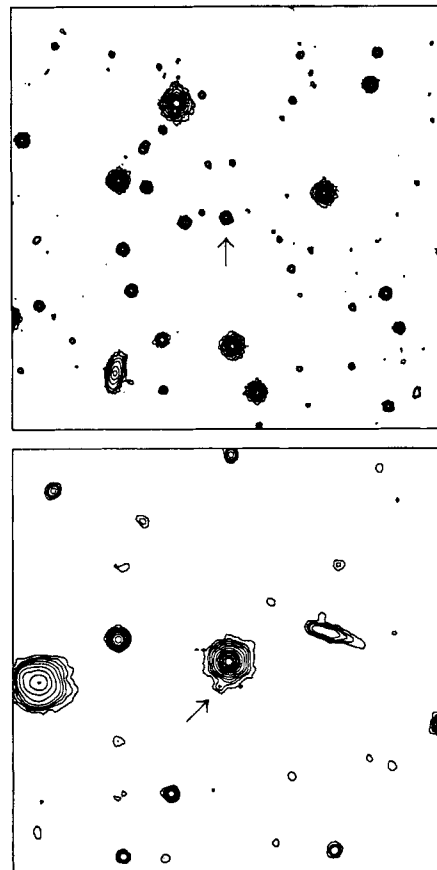


FIG. 1 Finding charts for the field. The top image is 3 arcmin on each side, from V- and R-band images obtained at the Palomar 60-inch telescope. The quasar is marked with the arrow; its coordinates are right ascension 22 h 33 min 51.15 s, declination $+13^\circ 10' 46''$ (B1950 equinox). The bottom image is 40 arcsec on each side, from R-band images obtained at the Keck telescope. The DLA galaxy is indicated with the arrow. In both images north is up and east to the left, and the intensity contours are spaced logarithmically.

difference of $\Delta V = 209 \text{ km s}^{-1}$ from the absorption system (Figs 2 and 3). The central wavelength of the emission line may be affected by absorption in the ambient gas, which could modify its observed redshift. The observed line flux is $F_{1216} = (6.4 \pm 1.2) \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$, where flux zero point uncertainties dominate the error. Although no other lines were detected, the close proximity to the known absorber redshift and the consistent colours and magnitudes leave little doubt as to the redshift identification. There are no high-ionization lines detected (for example, N v 1,240 $\text{\AA}$, or C iv 1,549 $\text{\AA}$, which would indicate the presence of an AGN) with 1σ upper limits of $\sim 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1}$.

Subsequent infrared observations were obtained at the Keck and the Kitt Peak 4-m telescope. Our measurements are consistent with the estimate¹⁷ of $K \approx 22 \pm 0.3$ mag. We have an upper limit to the [O III] 5,007 $\text{\AA}$ line flux of $F_{5007} < 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$, which is consistent with expectations based on our measurement of the Ly α line flux and simple photoionization models. This suggests that the object is not highly reddened.

To estimate the physical properties of the object, we assume a standard Friedman model cosmology with $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_0 = 0.2$ (our principal conclusions are not sensitive to the exact choice of these parameters). This gives the look-back time of 83% of the age of the Universe at this redshift. Assuming the onset of galaxy formation at $z \approx 5$, this object is only $\sim 7 \times 10^8$ yr old, which corresponds to only a few free-fall times for a normal galaxy.

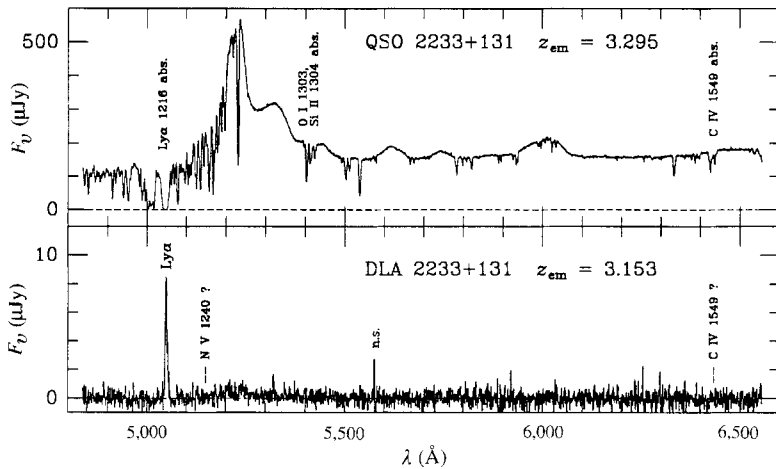


FIG. 2 Intermediate-resolution spectra of the quasar (QSO, top) and the damped Ly α galaxy (DLA, bottom), obtained at the Keck telescope. Absorption features due to the DLA galaxy are marked in the spectrum of the quasar; in addition, the Si II 1,260 Å line is present in absorption at $\lambda = 5,231$ Å, but it may be confused with the associated Ly α absorption in the quasar spectrum itself. Strong Ly α emission is seen in the spectrum of the DLA galaxy. Note the complete absence of high-ionization lines of N V 1,240 Å and C IV 1,549 Å (marked at the appropriate locations), suggesting that there is no detectable active nucleus in this object. The glitch at 5,577 Å is due to the poor night sky line subtraction. The continuum, but not the Ly α line, is oversubtracted in this spectrum, due to the presence of a much brighter quasar nearby; the true continuum level should be around 0.35 μ Jy.

This, and the very blue continuum implied by our photometry, suggests that DLA2233 + 131 must be young. For our assumed cosmology, the distance modulus is $(m - M) = 47.11$, the angular size distance (appropriate for converting observed angular separations into the rest-frame linear sizes) is 1.96×10^{28} cm, and the luminosity distance (appropriate for converting observed fluxes into the rest-frame luminosities) is 8.14×10^{28} cm.

At present we cannot be sure if the observed emission is coming

from a young galaxy's disk, or its bulge, or both. Deep Hubble Space Telescope images may help resolve this question. This uncertainty should be borne in mind in the discussion that follows.

The projected separation of the detected galaxy and the quasar (that is, its absorber portion) is then 17.2 kpc, comparable to the sizes of normal galaxy disks. The velocity field, if indeed due to rotation, is comparable to the rotation curve amplitudes of normal spirals, for reasonable projection angles. Moreover, the Ly α line shows a red wing (Fig. 3), as may be expected from lines of sight intersecting differentially rotating galaxy disks^{18,19}. The implied dynamical mass is

$$M_{\text{dyn}} \geq \frac{V^2 R}{G} = 1.86 \times 10^{11} M_{\odot} \left(\frac{V}{200 \text{ km s}^{-1}} \right)^2 \left(\frac{R}{20 \text{ kpc}} \right)$$

(a lower limit due to the unknown projection effects). From the observed column density of the neutral hydrogen, $N_{\text{HI}} = 1.0 \times 10^{20} \text{ cm}^{-2}$ (ref. 4), we can estimate the gas mass of

$$M_{\text{HI}} \approx \pi \langle R \rangle^2 N_{\text{HI}} m_p \approx 1.0 \times 10^9 M_{\odot} \left(\frac{R}{20 \text{ kpc}} \right)^2$$

where m_p is the mass of the proton. These values are typical of normal galaxy disks today. They are only meant to be indicative, given the uncertainties about the morphology of the object and the projection effects.

The implied Ly α line luminosity is $L_{1216} = 5.3 \times 10^{42} \text{ erg s}^{-1}$. Estimates of conversion of the Ly α line luminosity to the implied unobscured star formation rate (SFR) are in the range $L_{1216} = (7 \pm 4) \times 10^{41} \text{ erg s}^{-1}$ for an SFR of $1 M_{\odot} \text{ yr}^{-1}$, depending on the stellar initial mass function (IMF)^{20,21}. We thus estimate the unobscured SFR in DLA2233 + 131 to be $7.5^{(+10)}_{(-3)} M_{\odot} \text{ yr}^{-1}$. An independent estimate of the SFR can be obtained from the rest-frame continuum luminosity at 1,500 Å. For our assumed cosmology, $P_{1500} = 9.4 \times 10^{40} \text{ erg s}^{-1} \text{ Å}^{-1}$. Population synthesis models²² for a constant SFR of $1 M_{\odot} \text{ yr}^{-1}$, assuming a Salpeter IMF, predict $P_{1500} \approx 1.5 \times 10^{40} \text{ erg s}^{-1} \text{ Å}^{-1}$, but plausible variations in the IMF slope can change that number by a factor of 3. With this conversion, we derive an SFR of $\sim 6.4 M_{\odot} \text{ yr}^{-1}$, in good agreement with our estimate from the Ly α line.

This SFR, only a few times higher than in most spiral galaxies today, is an order of magnitude less than the observed star formation rates in ultraluminous IRAS galaxies, or expected rates in proto-ellipticals, but it is comparable to what may be expected from a gradually forming, young disk. This may be a lower limit, if some line emission is absorbed by the gas or dust. We derive the rest-frame equivalent width of 37 Å, less than what most models would predict^{20,21} for our estimated SFR, which can be explained by a very modest amount of absorption, or the effect of superposed stellar Ly α absorption.

Some of the observed Ly α emission may be due to the ionization of the H I cloud by the metagalactic ultraviolet flux,

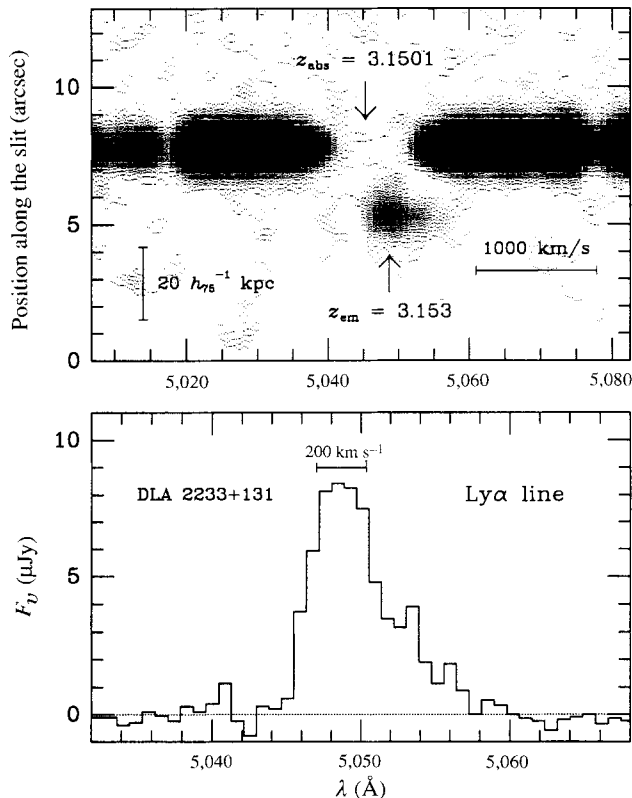


FIG. 3 The top panel shows a magnified view of the spectroscopic frame, showing the Ly α line emission from the DLA galaxy, and the associated absorption in the quasar spectrum. Their central redshifts are indicated, and the corresponding rest-frame velocity difference is only 209 km s^{-1} , comparable to rotation curve amplitudes of normal disk galaxies. The projected angular and rest-frame velocity scales are indicated where h_{75} is the Hubble constant in units of $75 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Note the red wing in the Ly α line, which is better seen in the bottom panel, showing a magnified region of the extracted spectrum. This is a characteristic line shape expected from differentially rotating galaxy disks.

which at this redshift is estimated to be in the range $J_v \approx (1.5 \pm 0.5) \times 10^{21} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}$ (ref. 23). Following Wolfe *et al.*⁵, at $z = 3.153$ this implies the observed Ly α surface brightness of $(1.8 \pm 0.6) \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ arcsec}^{-2}$, which is a negligible fraction of the observed line flux.

We can also extrapolate the observed continuum flux to the rest-frame B band. Assuming a flat spectrum, we obtain for the observed flux $F_v(B_{\text{rest}}) = 0.36 \mu\text{Jy}$, and for a power law spectrum $F_v \approx \nu^\alpha$, with $\alpha = -0.4$, the reddest spectrum compatible with our data, $F_v(B_{\text{rest}}) = 0.53 \mu\text{Jy}$. In the assumed cosmology, these correspond to the absolute magnitudes $M_B = -20.33$ and -20.75 , respectively. For comparison, an L^* galaxy (that is, having luminosity typical of a present-day galaxy) has $M_B = -20.4$ for the same cosmology. Thus, the optical luminosity of DLA2233 + 131 corresponds to that of normal, evolved disk galaxies today, and is at least two orders of magnitude lower than that of typical quasars at that redshift. Given its higher star formation rate, we conclude that it yet has to make most of its stars.

This discovery should then be considered in the context of searches for protogalaxies. Despite considerable efforts, no obvious population of progenitors of normal galaxies has been found in the past^{24–27}. One promising technique is to select high-redshift galaxies through their continuum colours, using the Lyman discontinuity at 912 Å (refs 16,17). Most searches to date have concentrated on emission-line objects, primarily the Ly α line redshifted in the optical window, or the hydrogen Balmer lines or nebular oxygen lines redshifted into the near-infrared^{20,24–29}. Many interesting high-redshift galaxies have been found in this way, but essentially all of them either contain AGN, or are located in close proximity to one, and thus may be powered by reprocessed AGN radiation, rather than star formation. It is still possible that many or all of the known high-redshift AGN are situated in young galaxies.

DLA systems represent an already well established, large population of high-redshift objects for which the confusion with AGN does not arise. They have been proposed as likely progenitors of normal disk galaxies¹, on the basis of their properties as inferred from the absorption signatures alone. In support of this idea, the properties of DLA2233 + 131, a field galaxy now clearly identified with a DLA system, are exactly what may be expected of a young, still-forming disk galaxy at a high redshift. This object may eventually evolve into a normal spiral galaxy not very different from our own. This is also fully consistent with the interpretation of Lyman-break objects at comparable redshifts as a population of protobulges¹⁷. The populations of DLA systems, of which DLA2233 + 131 may be representative, and of Lyman-break objects may overlap considerably, and they may be tentatively identified as progenitors of typical normal galaxies today. □

Received 1 February; accepted 21 May 1996.

- Wolfe, A. M. *Ann. N.Y. Acad. Sci.* **688**, 281–296 (1993).
- Lanzetta, K. M., Wolfe, A. M. & Turnshek, D. A. *Astrophys. J.* **440**, 435–457 (1995).
- Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J. & Hazard, C. *Astrophys. J.* **427**, L13–L16 (1994).
- Lu, L., Wolfe, A. M., Turnshek, D. A. & Lanzetta, K. M. *Astrophys. J. Suppl. Ser.* **84**, 1–38 (1993).
- Wolfe, A. M., Turnshek, D. A., Lanzetta, K. M. & Oke, J. B. *Astrophys. J.* **385**, 151–172 (1992).
- Elston, R., Bechtold, J., Lowenthal, J. D. & Rieke, M. *Astrophys. J.* **373**, L39–L42 (1991).
- Hu, E. M., Songaila, A., Cowie, L. L. & Hodapp, K.-W. *Astrophys. J.* **419**, L13–L16 (1993).
- Lowenthal, J. D. *et al. Astrophys. J.* **451**, 484–497 (1995).
- Lowenthal, J. D. *et al. Astrophys. J.* **377**, L73–L77 (1991).
- Macchetto, F., Lipari, S., Giallisco, M., Turnshek, D. A. & Sparks, W. B. *Astrophys. J.* **404**, 511–520 (1993).
- Francis, P. J. *et al. Astrophys. J.* **457**, 490–499 (1996).
- Warren, S. J. & Möller, P. *Astr. Astrophys.* **311**, 25–35 (1996).
- Möller, P. & Warren, S. J. *Astr. Astrophys.* (in the press).
- Crampton, D., Schade, D. & Cowley, A. P. *Astr. J.* **90**, 987–997 (1985).
- Oke, J. B. *et al. Publ. astr. Soc. Pacif.* **107**, 375–385 (1995).
- Steidel, C. C., Pettini, M. & Hamilton, D. *Astr. J.* **110**, 2519–2536 (1995).
- Steidel, C. C., Giallisco, M., Pettini, M., Dickinson, M. & Adelberger, K. *Astrophys. J.* **462**, L17–L21 (1996).
- Lanzetta, K. M. & Bowen, D. V. *Astrophys. J.* **391**, 48–72 (1992).
- Wolfe, A. M. *et al. Astrophys. J.* **435**, L101–L104 (1994).
- Thompson, D., Djorgovski, S. & Trauger, J. *Astr. J.* **110**, 963–981 (1995).
- Charlot, S. & Fall, S. M. *Astrophys. J.* **415**, 580–588 (1993).
- Leitherer, C., Robert, C. & Heckman, T. M. *Astrophys. J. Suppl. Ser.* **99**, 173–187 (1995).
- Haardt, F. & Madau, P. *Astrophys. J.* **461**, 20–37 (1996).

- Pritchett, C. J. *Publ. astr. Soc. Pacif.* **106**, 1052–1067 (1994).
- Djorgovski, S. & Thompson, D. in *The Stellar Populations of Galaxies* (eds Barbu, B. & Renzini, A.) 337–347 (IAU Symp. 149, Kluwer, Dordrecht, 1992).
- Djorgovski, S. in *New Light on Galaxy Evolution* (eds Bender, R. & Davies, R.) 277 (IAU Symp. 171, Kluwer, Dordrecht, 1996).
- Djorgovski, S. in *Cosmology and Large-Scale Structure in the Universe* (ed. de Carvalho, R.) 73–95 (Conf. Ser. 24, Astr. Soc. Pacif., San Francisco, 1992).
- Thompson, D. & Djorgovski, S. *Astr. J.* **110**, 982–993 (1995).
- Pahre, M. A. & Djorgovski, S. *Astrophys. J.* **449**, L1–L4 (1995).

ACKNOWLEDGEMENTS. This work was based in part on the observations obtained at the W. M. Keck Observatory, which is operated jointly by the California Institute of Technology and the University of California. We thank the staff of Keck and Palomar observatories for their help during our observing runs, and C. Steidel, L. Lu and M. Rauch for useful conversations. S.G.D. was supported by the US National Science Foundation, and the Bressler Foundation.

CORRESPONDENCE should be addressed to S.G.D. (e-mail: george@deimos.caltech.edu).

Microscopic coexistence of magnetism and superconductivity in $\text{ErNi}_2\text{B}_2\text{C}$

U. Yaron*, P. L. Gammel*, A. P. Ramirez*, D. A. Huse*, D. J. Bishop, A. I. Goldman†, C. Stassis†, P. C. Canfield‡, K. Mortensen‡ & M. R. Eskildsen‡

* Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

† Ames Laboratory and Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA

‡ Risø National Laboratory, 4000 Roskilde, Denmark

MAGNETISM and superconductivity are manifestations of two different ordered states into which metals can condense at low temperatures. In general these states are mutually exclusive¹; they do not coexist at the same place in a sample. The study of the interplay between these properties has recently been revitalized by the discovery^{2,3} of a class of compounds with formula $\text{RNi}_2\text{B}_2\text{C}$ (where R is a rare-earth element) which are both antiferromagnetic and superconducting at sufficiently low temperature⁴. It has been suggested⁵ that magnetic and superconducting order can coexist in these materials on an atomic scale. Here we use small-angle neutron scattering to study the structure of the superconducting vortex lattice in $\text{ErNi}_2\text{B}_2\text{C}$. Our results show that the development of magnetic order causes the vortex lines to disorder and rotate away from the direction of the applied magnetic field. This coupling of superconductivity and magnetism provides clear evidence for microscopic coexistence of magnetic and superconducting order, and indicates that magnetic superconductors may exhibit a range of unusual phenomena not observed in conventional superconductors.

Our study has found a number of novel results. First, we find that the vortex structure is a square lattice as opposed to an hexagonal one as is more typically found in a high-field, strongly type II superconductor. For this material⁶ the Ginzburg–Landau parameter is intermediate between the high values found for the high- T_c superconductors and low values for marginally type II systems such as Nb. In addition, we find a significant coupling between the magnetic ordering and the flux lines as evidenced by a rotation of the vortex lines away from the direction of the applied field below the Néel transition as well as a disordering of the lattice. We believe that these represent the first data to show how developing magnetic order can directly influence the vortex lattice.

Our experiments were performed on a large single crystal of $\text{ErNi}_2\text{B}_2\text{C}$ ($6 \times 6 \times 2 \text{ mm}$) grown using an unusual high-temperature flux method with Ni_2B as a solvent⁶. In order to allow neutrons to penetrate the sample, isotopically enriched B^{11} was used. The

sample was characterized by X-ray, transport, and longitudinal and transverse magnetization measurements. Shown in Fig. 3 inset is the phase diagram for this system for the field orientation we have used ($\mathbf{H} \parallel \mathbf{c}$). The zero-field T_c of this compound is 10.5 K, and the Néel transition occurs at $T_N = 6$ K. As shown in the phase diagram, there exist four phases in the H - T plane—a paramagnetic normal metal above the upper critical field, H_{c2} , and above T_N , an antiferromagnetic normal metal above H_{c2} and below T_N , an antiferromagnetic superconductor at temperatures below T_N and fields below H_{c2} , and a paramagnetic, superconducting state for temperatures above T_N and fields below H_{c2} . The strong interaction between the magnetic order of the Er spins and the superconductivity is evidenced by the change in slope and magnitude of $H_{c2}(T)$ at T_N . We have imaged the Abrikosov vortex lattice using small-angle neutron scattering (SANS) at the cold-neutron guide hall at Risø National Laboratory. In the experiment, a superconducting magnet applies the field nearly parallel to both the c -axis of the crystal as well as the incident neutron beam. The neutrons, which are Bragg-scattered from the magnetic field pattern due to the vortex lines in the sample, are counted by an area detector at the end of a 6-m evacuated chamber.

Figure 1 shows the results of a typical experiment; it is a grey-scale image of the reflectivity on the plane of the detector for an experiment where the field is applied roughly parallel to the c -axis and at a temperature such that the sample is in the magnetic, superconducting state. The most obvious feature of the data shown is the square symmetry of the lattice. This square lattice is found over the whole range of our neutron experiment ($2 \text{ kOe} < H < 12 \text{ kOe}$; $1.6 \text{ K} < T < 8.0 \text{ K}$). In addition to the bright (10) peaks, we can also see the higher-order (11) peaks. Measurements of the magnetic order (Fig. 1 inset) show that the

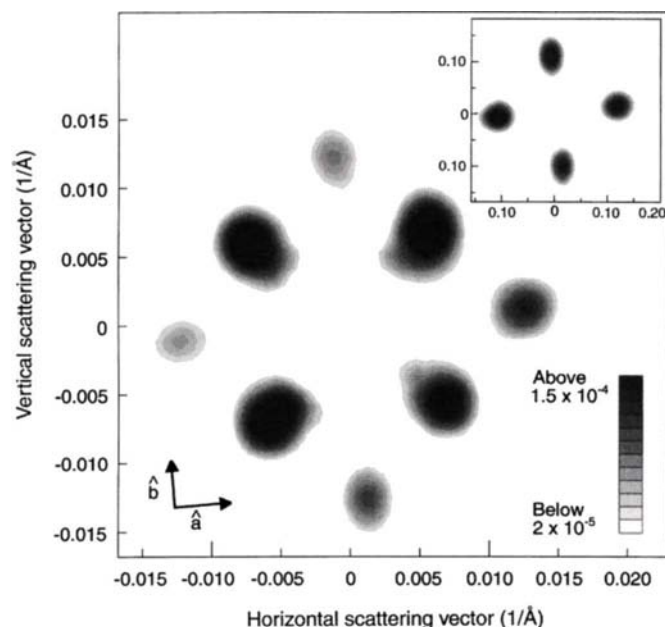


FIG. 1 Grey-scale image of the neutron reflectivity for the flux-lattice Bragg scattering ($H_c = 4 \text{ kOe}$, $T = 1.6 \text{ K}$). The rocking curves are sufficiently broad in comparison to the scattering angle that all eight peaks are illuminated when the incident neutron beam is parallel to $\mathbf{H}$. For scattering from the vortex lattice, we have used neutron wavelengths of $\lambda_n \approx 9.0 \text{ Å}$, $\Delta\lambda_n/\lambda_n \approx 18\%$ and incident beam divergences of 0.15° FWHM in both the horizontal and vertical directions. The inset shows the 11th-order peaks from the magnetic scattering, below T_N at $T = 1.6 \text{ K}$, $H = 0$. By moving the area detector to within one metre of the cryostat and by going to much shorter-wavelength neutrons ($\lambda_n \approx 2.8 \text{ Å}$), we can image the magnetic order of the Er spins below the Néel transition at 6 K . Thus, without changing the experimental apparatus, we can accurately determine both the crystal orientation as well as the flux-lattice orientation.

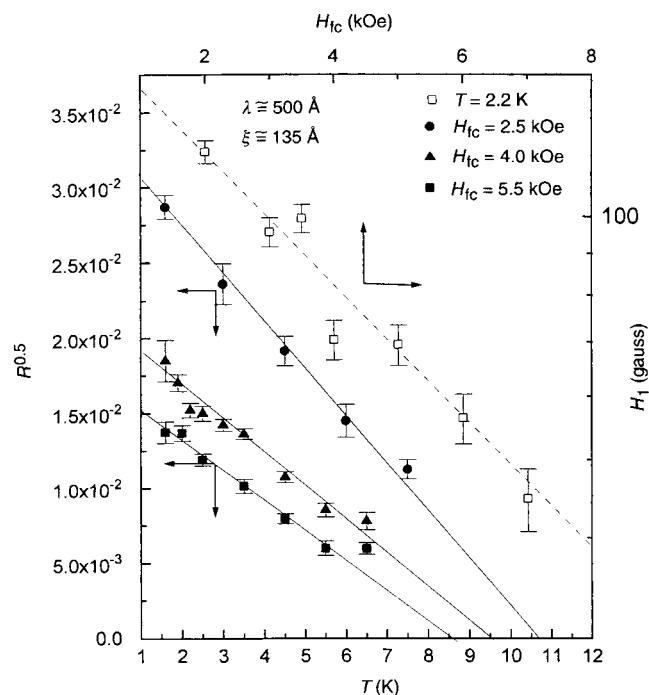


FIG. 2 Plot of the log of the form factor H_1 versus field at 2.2 K , and the square root of the reflectivity R versus temperature for various fields.

flux lattice is rotated 45° with respect to the underlying crystal lattice. The magnetic structure^{7,8} in this material is a transverse, incommensurately modulated spin wave with wavevector $0.553 a^*$. Shown in the inset are the 11th-order peaks. The brightness of these high-order reflections shows that the spin wave in this material has an essentially rectangular modulation. There are four such peaks because at the Néel transition domains form at right angles to each other, there being two equivalent directions for the spin wave in this tetragonal system.

Fourfold symmetries for a flux-line lattice (FLL) have been seen for marginally type II superconductors such as Nb. In those materials the vortex lattice is strongly influenced by gap and Fermi surface anisotropies. The vortex lattice structure shown in Fig. 1 also has a casual similarity to that found in heavily-twinned $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ (YBCO) samples but is different in an important detail^{9,10}. In YBCO there are four first-order (10) peaks, but no higher-order (11) peaks; this is understood to be the result of the twins pinning the vortex lines along their length. Thus the pattern seen in YBCO is that of two, one-dimensional structures at right angles to each other and superimposed upon each other. This is to be contrasted with the data shown in Fig. 1 in which there are higher-order peaks. An analysis of the intensity of these higher-order peaks indicates that this is a true square vortex lattice that has moderately long-ranged order with $R_c/a_0 > 5$, where a_0 is the inter-vortex distance and R_c is the positional correlation length in the plane. The origin of this square vortex lattice is an open question, but there has been speculation that a flux lattice of this structure and orientation could be stabilized in strongly type II superconductors owing to magneto-elastic coupling¹¹.

In Figure 2 we show the field- and temperature-dependence of the scattered intensity in the lowest-order vortex lattice Bragg peaks. From the field dependence of the intensity we can extract the penetration depth and coherence length^{12,13}. The neutron reflectivity can be expressed as $R = (2\pi\gamma_n^2 t / 16\phi_0^2 q) H_1^2$ where $\gamma_n = 1.91$ is the neutron gyromagnetic ratio, t is the sample thickness and $\phi_0 = hc/2e$ is the flux quantum. The form factor H_1 has the units of a magnetic field and is given by $H_1 = (\phi_0/\lambda^2)(\frac{1}{4}\pi^2) \exp(-2\pi^2 B_c^2/\phi_0)$ for B well below H_{c2} . Therefore,

a plot of the log of the form factor versus the magnetic field should be a straight line whose slope is determined by the coherence length ξ and whose intercept at zero field is related to the penetration depth λ . Such an analysis of the data shown in Fig. 2 indicates that $\lambda \approx 500 \text{ \AA}$ and $\xi \approx 135 \text{ \AA}$ at $T = 2.2 \text{ K}$ (ref. 6). The data in Fig. 2 also shows how the intensity varies with temperature. Assuming a mean-field temperature dependence $\lambda \sim (1 - T/T_c(H))^{-1/2}$, we would expect the square root of the reflectivity R to scale linearly with temperature going to zero at $T_c(H)$. Shown are the data for 2.5, 4.0 and 5.5 kOe which follow this trend reasonably well. An increase over this at $T_N = 6 \text{ K}$ is expected from the increased $H_{c2} = \phi_0/(2\pi\xi^2(T))$. Using published H_{c2} values⁶ would actually lead to an increase roughly three times the observed value, although this could be compensated for by an increase in λ . Clearly, more experimental and theoretical work needs to be done to elucidate the changes in the microscopic superconducting parameters due to spin ordering at T_N .

Shown in Fig. 3a is the temperature- and field-dependence of the full width at half maximum (FWHM) of the rocking curve, σ_m . If the width of the rocking curve is due to microscopic disorder in the flux lines, then σ_m is proportional to the inverse of the longitudinal correlation length ξ_L —the distance over which the vortex lines remain correlated to each other along their length, parallel to the applied field. A rocking curve is a plot of the scattered intensity in the Bragg peaks as the sample and the magnet are rotated through the Bragg condition around an axis perpendicular to both the incoming neutron beam and the scattering wavevector. The temperature- and field-dependence of σ_m shows a number of unusual features. First, it is roughly field-independent over our measured field range. Second, it increases sharply at low temperatures. Both of these results are in stark contrast to what is found in conventional, type II superconductors. In previous

studies¹³, the straightness of the flux lines was related to the transport critical current as $\sigma_m \approx (J_c/B^2\xi)^{1/2}$ where J_c is the critical current. This predicts a field dependence which should decrease σ_m by ~ 2 with fixed J_c . The expected change should be even greater (~ 4) as J_c is known to decrease with increasing field. Clearly σ_m does not follow this model. The magnitude of σ_m is also quite large given the small mosaic spread of the crystal ($\sim 0.05^\circ$) and small critical current ($J_c \approx 100 \text{ A cm}^{-2}$). This suggests that the critical current and the rocking-curve width probe different length scales in the sample. One possibility is that the rocking-curve width is not due to the microscopic pinning-induced disorder in the vortex lattice, but instead represents variations in the orientation of the vortex lattice at longer length scales. The temperature dependence is also quite surprising. The temperature dependence of J_c and ξ would predict very little temperature dependence of σ_m below $\sim 0.5 T_c$. We see a pronounced rise at low temperatures, suggesting that the development of the magnetic order disorders the lattice as the temperature is decreased. Thus, this might be a system in which the lattice could be transformed into a glass by cooling it—a most surprising possibility.

In Fig. 3b is plotted the measured angle between the applied magnetic field H_a and the flux lines in the sample. This angle is measured by determining the locations of the peaks of the rocking curves. At high temperatures, as we would expect, we find that the vortex lattice is well aligned with the applied field. However, as the sample is cooled, an unexpected thing happens—the vortex lattice rotates away from the applied field direction in the direction away from the c -axis. In our experiment, by studying the antiferromagnetic peaks we know that the magnetic field is applied at an angle of 1.6° away from the exact c -axis of the sample in the a - c plane of the sample. At relatively high temperatures, between T_c and $\sim 4 \text{ K}$, the average orientation of the vortex lattice is well aligned with H_a . But at roughly 4 K it begins to rotate away from the applied field direction, towards the a -direction but still in the a - c plane. This rotation can be as much as 1° . Measurements of the transverse magnetization quantitatively agree with the SANS results, and confirm that the average internal field in the sample rotates relative to the applied field. The transverse magnetization is in accord with the recent report¹⁴ of weak ferrimagnetism in this system for $T < 2.3 \text{ K}$. The internal ferrimagnetic field can then add to the externally applied field and rotate it. The development of low-temperature magnetic order in this system therefore both disorders and rotates the flux lattice.

We believe that these results are the first to demonstrate how, in a magnetic superconductor, there can be a direct coupling between the magnetic-order parameter and the vortex lattice structure. The data we have shown here relating the vortex lattice structure and the magnetic order suggests that there will be a rich and interesting phenomenology in these systems to be studied. In particular, the unusual static properties of the lattice suggests that there will be unconventional and interesting dynamics in this and other magnetic superconductors. $\square$

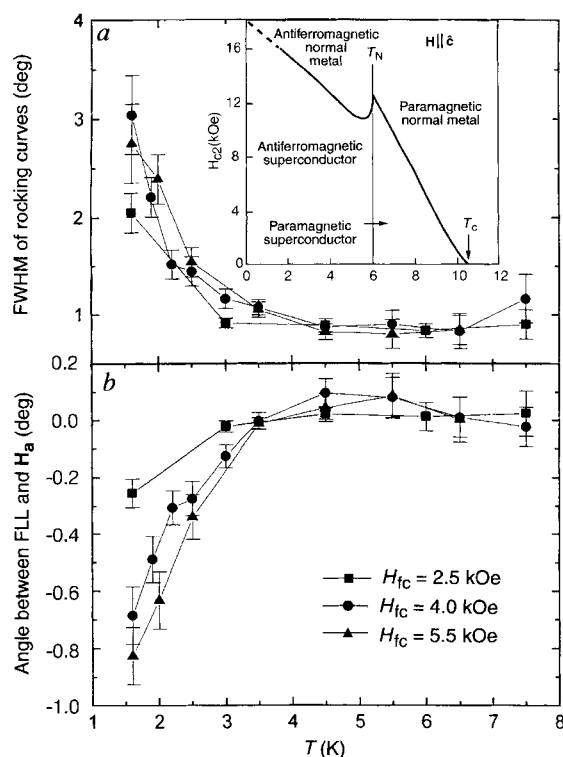


FIG. 3 Plot of the FWHM of the rocking curves as a function of temperature for various fields (a) and the angle between the flux-line lattice (FLL) and the applied field as a function of temperature for various fields (b). The instrumental resolution is roughly an order of magnitude smaller than the FWHM, and we can make the approximation that $\text{FWHM} = \sigma_m$. The inset in (a) shows the phase diagram for $\text{ErNi}_2\text{B}_2\text{C}$ (after ref. 6).

Received 22 March; accepted 11 June 1996.

1. Matsubara, T. & Kotani, A. (eds) *Superconductivity in Magnetic and Exotic Materials* (Springer, Berlin, 1984).
2. Nagarajan, R. et al. *Phys. Rev. Lett.* **72**, 274–277 (1994).
3. Cava, R. J. et al. *Nature* **367**, 252–253 (1994).
4. Cho, B. K., Canfield, P. C. & Johnston, D. C. *Phys. Rev.* **B52**, R3844–R3847 (1995).
5. Eisaki, H. et al. *Phys. Rev.* **B50**, 647–650 (1994).
6. Cho, B. K. et al. *Phys. Rev.* **B52**, 3684–3695 (1995).
7. Zarestky, J. et al. *Phys. Rev.* **B51**, 678–680 (1995).
8. Sinha, S. K. et al. *Phys. Rev.* **B51**, 681–684 (1995).
9. Cubitt, R. et al. *Nature* **365**, 407–411 (1993).
10. Keimer, B. et al. *Science* **262**, 83–86 (1993).
11. Ullmaier, H., Zeller, R. & Dederichs, P. H. *Phys. Lett.* **44A**, 331–332 (1973).
12. Kleiman, R. N. et al. *Phys. Rev. Lett.* **69**, 3120–3124 (1992).
13. Gammel, P. L. et al. *Phys. Rev. Lett.* **72**, 278–282 (1994).
14. Canfield, P. C., Bud'ko, S. L. & Cho, B. K. *Physica C* **262**, 249–254 (1996).

ACKNOWLEDGEMENTS. We thank P. Littlewood, B. Batlogg and C. Varma for many helpful discussions; C. Broholm and T. Siegrist for technical assistance in the initial characterization of the samples; and J. Lebeck, B. Breiting and the staff at Risø for their assistance.

CORRESPONDENCE should be addressed to D.J.B. (e-mail: djb@lucent.com).

Entropically driven binding in a self-assembling molecular capsule

Jongmin Kang & Julius Rebek Jr

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE encapsulation of molecules within larger molecular or supramolecular cages (see ref. 1 for examples) brings about many interesting phenomena, among them the stabilization of reactive intermediates², new forms of stereoisomerism³, and templating of the cage itself^{4,5}. Here we describe a self-assembling hydrogen-bonded molecular capsule that encapsulates guest molecules in a reversible, apparently entropy-driven process, giving rise to unusual temperature dependence of the formation process. The positive entropy of formation—at first sight unexpected—seems to be the result of the fact that more than one solvent molecule is included in the 'bare' cage, making their replacement by a single large molecule (here adamantane or ferrocene) entropically favourable.

Structures **1** (Fig. 1) were recently introduced as self-assembling capsules capable of reversibly binding complementary molecules in solution⁶. These compounds dimerize through hydrogen bonding when the concave surfaces of two molecules of **1** are brought together. The dimerization gives a closed-shell capsule of roughly spherical shape, resembling a softball held together with eight hydrogen bonds along its seam (Fig. 1). The information for self-assembly is not unique to dimerization, because higher-order aggregates, probably chain-like oligomers, also emerged in solutions of **1**. To direct the assembly towards capsules, structure **2** has now been prepared. It features additional hydrogen-bonding sites provided by the phenolic functions. These sites are picketed along the seam in a way that stabilizes the capsule, and dimerization is favoured over other possible assemblies⁷.

Experimental procedures for the synthesis and characterization of **2** are outlined in supplementary information. The ¹H-NMR spectra in solvents such as *p*-xylene-*d*₁₀ (fully deuterated *p*-xylene) and CDCl₃ show sharp signals characteristic of an ordered, well defined assembly (Fig. 2*a, d*). The downfield shifts observed for

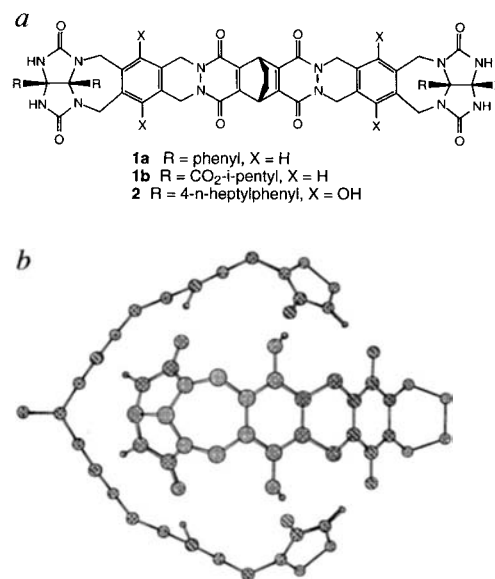


FIG. 1 *a*, Structures of the self-complementary molecules **1** and **2**. The glycoluril carbonyls at the ends of the molecule are the complementary hydrogen-bonding sites for the four phenolic donors of **2**. As a result, up to eight additional hydrogen bonds can stabilize the dimeric form of **2**. *b*, The hydrogen-bonding network of dimeric **2**, as calculated using MacroModel¹⁷ 3.5× with the AMBER force field and GB/SA chloroform solvation, is highlighted. The view is a cross-section of the capsule and other atoms have been deleted for clarity.

the glycoluril N–H (7.74 p.p.m.) and O–H (8.00 p.p.m.) resonances are concentration-independent and indicate extensive hydrogen bonding. The NMR spectra indicate that the dominant form of **2** in these solvents is the dimeric capsule, held together by a network of up to 16 hydrogen bonds. The shape and volume of the capsule is determined by the average geometries of all the hydrogen bonds involved, which is a compromise for the N–H and O–H donors. We present evidence that this capsule is robust and maintains its structure with a variety of guests, including solvent molecules.

Upon addition of adamantane derivatives or ferrocene deriva-

TABLE 1 Thermodynamic parameters of guest-inclusion process of the dimer **2** in different solvents

Guests	K_{298} (CDCl ₃)	ΔG (kcal mol ⁻¹)	ΔH (kcal mol ⁻¹)	ΔS (e.u.)
Adamantane	$4.4 \times 10 \pm 2.2$	-2.2 ± 0.1	5.6 ± 0.3	26.2 ± 1.3
1-Adamantanecarboxamide	$2.8 \times 10^3 \pm 1.4 \times 10^2$	-4.7 ± 0.2	4.9 ± 0.3	32.1 ± 1.6
1-Ferrocenemethanol	$3.8 \times 10^3 \pm 1.9 \times 10^2$	-4.9 ± 0.2	4.5 ± 0.2	31.6 ± 1.6
1-Adamantanecarboxylic acid	$5.6 \times 10^3 \pm 2.8 \times 10^2$	-5.1 ± 0.3	4.1 ± 0.2	31.1 ± 1.6
1-Ferrocenecarboxylic acid	$1.7 \times 10^4 \pm 8.5 \times 10^2$	-5.8 ± 0.3	2.5 ± 0.1	27.9 ± 1.4
1-Adamantanemethanol	$2.7 \times 10^4 \pm 1.4 \times 10^3$	-6.0 ± 0.4	7.8 ± 0.4	46.6 ± 2.3
K_{298} (<i>p</i> -xylene- <i>d</i> ₁₀)				
Adamantane	$1.7 \times 10^3 \pm 90$	-4.6 ± 0.2	2.2 ± 0.1	22.6 ± 1.1
Ferrocene	$3.3 \times 10^3 \pm 170$	-5.0 ± 0.3	2.3 ± 0.1	24.4 ± 1.2
K_{298} (benzene- <i>d</i> ₆)				
1-Adamantanecarboxylic acid	$2.3 \times 10 \pm 1.2$	-1.8 ± 0.1	3.2 ± 0.2	16.9 ± 0.8
1-Ferrocenecarboxylic acid	$5.3 \times 10 \pm 2.7$	-2.4 ± 0.1	10.9 ± 0.5	44.7 ± 2.2

K , association constant; ΔG , free energy; ΔH , enthalpy; ΔS , entropy. Association constants are calculated from the following equation, in which to disregard values for solvent (S) concentration is consistent with the common practice for cyclodextrins and other receptors:

$$[H \cdot S \cdot H] + G \xrightleftharpoons{K} [H \cdot G \cdot H]$$

where $K = [H \cdot G \cdot H] / [H \cdot S \cdot H][G]$.

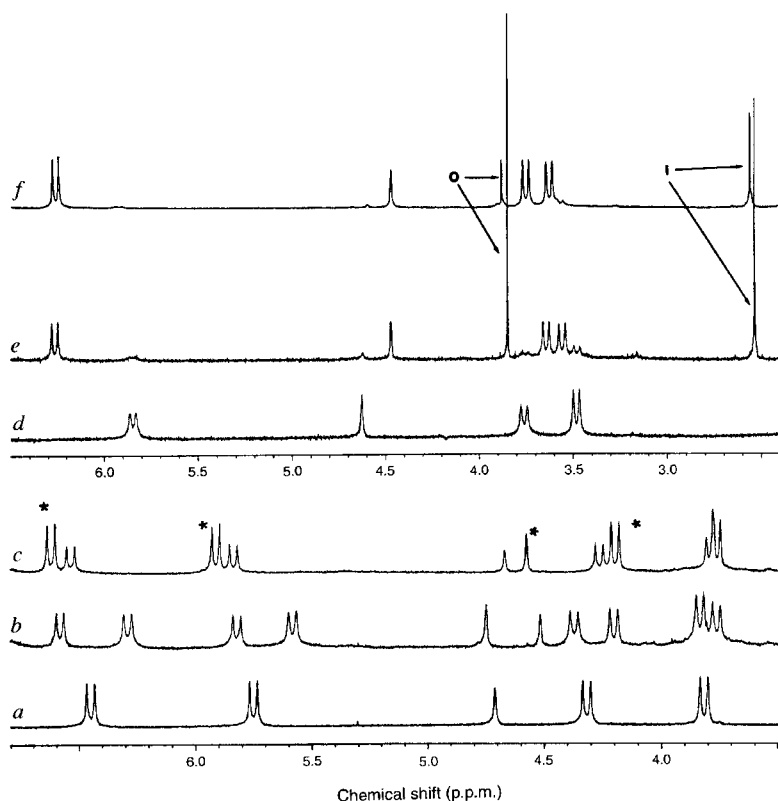


FIG. 2 ^1H NMR spectra of **2** and its encapsulation complexes in CDCl_3 and p -xylene- d_{10} . Signals of dimeric **2** with encapsulated guests are labelled with an asterisk. Guests inside and outside are labelled 'I' and 'O' respectively. a, Compound **2** alone at 298 K in CDCl_3 ; b, compound **2** with 0.7 equivalents of 1-adamantanecarboxylic acid added at 273 K in CDCl_3 ; c, compound **2** with 0.7 equivalents of 1-adamantanecarboxylic acid added at 323 K in CDCl_3 ; d, compound **2** alone at 298 K in p -xylene- d_{10} ; e, compound **2** with 1.2 equivalents of ferrocene added at 298 K in p -xylene- d_{10} ; f, compound **2** with 1.2 equivalents of ferrocene added at 343 K in p -xylene- d_{10} .

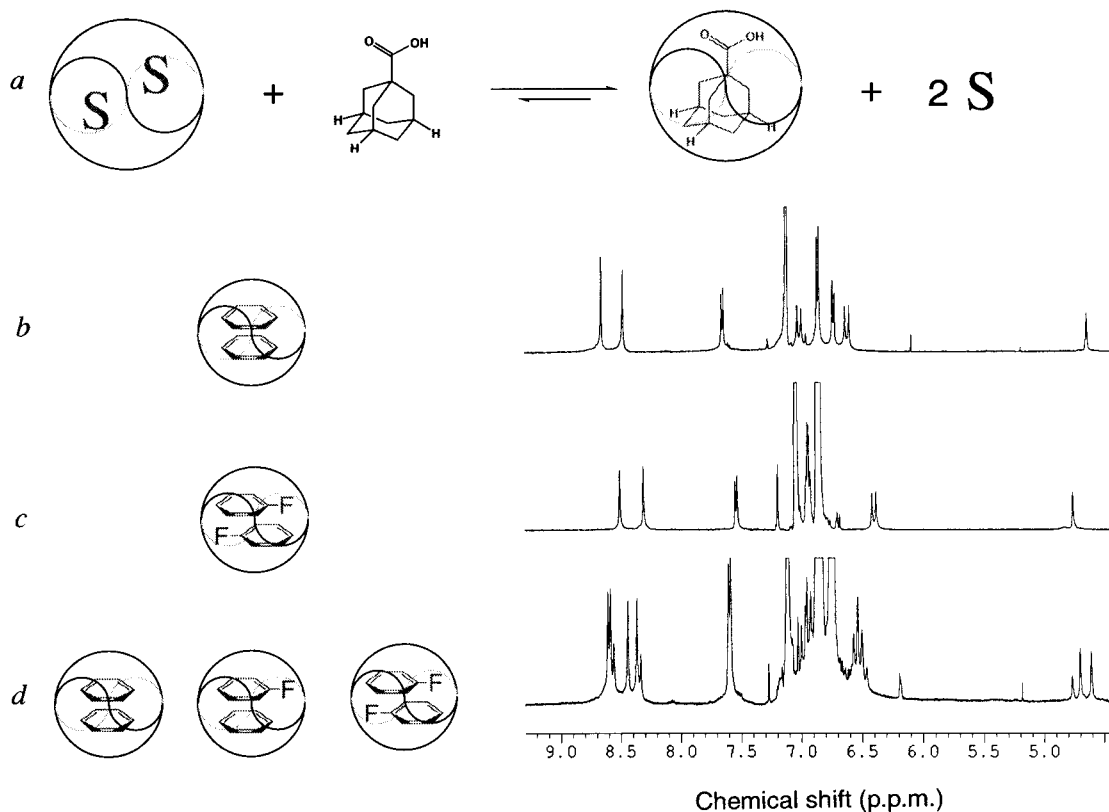


FIG. 3 a, Encapsulation of guests is driven by the liberation of more than one solvent molecule, S. b–d, ^1H NMR spectra of **2** in benzene- d_6 (b), in fluorobenzene- d_5 (c) and in a mixture of two parts benzene- d_6 and one part fluorobenzene- d_5 solvents (d). In the former two cases both guest molecules are equivalent; in the latter, there are three possible combinations, evident from the three signals at 4.5–4.8 p.p.m. and the three sets of signals at 8.3–8.7 p.p.m.

tives to a solution of **2** in CDCl_3 , benzene- d_6 or p -xylene- d_{10} , a second set of signals appears in the ^1H NMR spectra: a new capsule species is formed (Fig. 2b, e). In addition to the second set of signals for the capsule (host), a new set of upfield signals for the encapsulated guests emerges. These appear in the region 0.5 to -0.5 p.p.m. for the adamantanes, and 3.5 to 2.5 p.p.m. for the ferrocenes (additional spectra are available as supplementary information). Integration shows that the ratio of the new hosts to the encapsulated guests is 1:1. Moreover, the widely separated and relatively sharp signals for free and encapsulated guest in the NMR spectra indicate that exchange of guests in and out of the capsule is slow on the NMR scale. The relative concentrations of the original (solvent-occupied) and new (guest-occupied) capsules can be determined by signal integration. These values can be obtained for various guests as a function of temperature (Fig. 2c, f); equilibrium constants and their thermodynamic parameters could be determined from this procedure (Table 1).

The most significant observation is that guest encapsulation increases with temperature. As most processes involving host-guest association are entropically unfavourable but enthalpically favourable, the trends for **2** are unexpected. Indeed, much has been written about the compensating effects of entropy and enthalpy in complex formation⁸⁻¹⁰, but for **2** the inclusion of guests involves an enthalpic cost, which is compensated by a larger entropic gain. This behaviour is reminiscent of the classical hydrophobic effect, in which release of bound water to the bulk solvent compensates for the association of solutes. In organic media, such behaviour is rarely encountered and generally unpredictable¹¹⁻¹³, even though liberation of solvent is a universal feature of molecular recognition phenomena in solution. What is special about the case at hand?

It is proposed that for **2**, a single solvent molecule is too small to fill the cavity. Rather, more than one CDCl_3 or p -xylene is required to maximize the intermolecular forces (van der Waals interactions) between the convex surfaces of the guests and the concave surface of the host's interior. Now single molecular guests that fill the cavity and offer chemical and structural complementarity are entropically preferred to multiple solvent molecules; one guest releases several solvent molecules. Consequently, the encapsulation of adamantanes or ferrocenes by **2** increases entropy, because more than one molecule of solvent is released to the bulk solution (Fig. 3a). Release of solvents on the surface of the guest as encapsulation proceeds will reinforce the observed increases in enthalpy and entropy. Spectra taken in mixtures of two solvents ($\text{CDCl}_3/\text{CD}_2\text{Cl}_2$ and toluene- d_8/p -xylene- d_{10}) showed only a single (rather than two¹⁴) capsule species. Apparently, the spectra of the variously occupied capsules in these solvents are not resolved at the 500 MHz of the spectrometer. However, in mixtures of fully deuterated benzene and fluorobenzene, three species of capsule can be observed, a result that confirms the occupancy of the capsule with two solvent molecules (Fig. 3b-d).

The ability of **2** to host more than one guest suggests that its cavity is roomy enough to act as a reaction chamber for bimolecular reactions. It will be interesting to determine whether catalysis of reactions with transition states of the appropriate size shape and functionality can be observed within the capsule^{15,16}. □

Received 21 December 1995; accepted 29 May 1996.

1. Cram, D. J. *Nature* **356**, 29-36 (1992).
2. Cram, D. J. & Tanner, M. E. *Angew. Chem. (int. edn Engl.)* **30**, 1024-1027 (1991).
3. Timmerman, P. et al. *Angew. Chem. (int. edn Engl.)* **33**, 2345-2348 (1994).
4. Chapman, R. G. & Sherman, J. C. *J. Am. chem. Soc.* **117**, 9081-9082 (1995).
5. Chapman, R. G., Chopra, N., Cochien, E. D. & Sherman, J. C. *J. Am. chem. Soc.* **116**, 369-370 (1994).
6. Meissner, R. S., Rebek, J. Jr & Mendoza, J. D. *Science* **270**, 1485-1488 (1995).
7. Lindsey, J. S. *New J. Chem.* **15**, 153-180 (1991).
8. Dunitz, J. D. *Chem. Biol.* **2**, 709-712 (1995).
9. Peterson, B. R., Wallimann, P., Carcanague, D. R. & Diederich, F. *Tetrahedron* **51**, 401-421 (1995).
10. Searle, M. S., Westwell, M. S. & Williams, D. H. *J. chem. Soc. Perkin Trans. 2*, 141-151 (1995).
11. Cancelli, J. et al. *Angew. Chem. (int. edn Engl.)* **28**, 1246-1248 (1989).
12. Cram, D. J., Choi, H.-J., Bryant, J. A. & Knobler, C. B. *J. Am. chem. Soc.* **114**, 7748-7765 (1992).

13. Cram, D. J., Blanda, M. T., Paek, K. & Knobler, C. B. *J. Am. chem. Soc.* **114**, 7765-7773 (1992).
14. Shimizu, K. D. & Rebek, J. Jr *Proc. natn. Acad. Sci. U.S.A.* **92**, 12403-12407 (1995).
15. Mock, W. L., Irra, T. A., Wepsiec, J. P. & Adhya, M. J. *org. Chem.* **54**, 5302-5308 (1989).
16. Walter, C. J., Anderson, H. L. & Sanders, J. K. M. *J. chem. Soc. chem. Commun.* 458-460 (1993).
17. Mohamadi, F. et al. *J. Comput. Chem.* **11**, 440-467 (1990).

ACKNOWLEDGEMENTS. We are grateful for advice and experimental help from R. Meissner, and for financial support from the NIH.

CORRESPONDENCE and requests for materials should be addressed to J.R. at The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, California 92037, USA.

Deglacial sea-level record from Tahiti corals and the timing of global meltwater discharge

Edouard Bard^{*†}, Bruno Hamelin^{*}, Maurice Arnold[‡], Lucien Montaggioni[§], Guy Cabioch^{||}, Gérard Faure[¶] & Francis Rougerie[#]

^{*} CEREGE, Université d'Aix-Marseille III, URA CNRS 132, Europôle de l'Arbois, BP 80, Aix-en-Provence cedex 4, France

[†] Institut Universitaire de France

[‡] Centre des Faibles Radioactivités, CNRS-CEA, 91198 Gif-sur-Yvette cedex, France

[§] Centre de Sédimentologie et de Paléontologie, Université de Provence-St-Charles, 3 Pl. V. Hugo, 13331 Marseille cedex 3, France

^{||} ORSTOM-TOA, UMR CNRS-ORSTOM GéoAzur (EP 125), BP 48, 06230 Villefranche-sur-Mer, France

[¶] Laboratoire d'Hydrobiologie Marine, Université de Montpellier II, case 096, Pl. E. Bataillon, 34095 Montpellier cedex 5, France

[#] ORSTOM Centre de Papeete, Papeete, Tahiti, French Polynesia

The timing of the last deglaciation is important to our understanding of the dynamics of large ice sheets¹ and their effects on the Earth's surface^{2,3}. Moreover, the disappearance of the glacial ice sheets was responsible for dramatic increases in freshwater fluxes to the oceans, which probably disturbed the ocean's thermohaline circulation and, hence, global climate⁴⁻⁷. Sea-level increases bear witness to the melting of continental ice sheets, but only two such records—from Barbados^{8,9} and New Guinea^{10,11} corals—have been accurately dated. But these corals overlie active subduction zones, where tectonic movements are large and often discontinuous (especially in New Guinea), so the apparent sea-level records may be contaminated by a complex tectonic component. Here we date fossil corals from Tahiti, which is far from plate boundaries (and thus is likely to be tectonically relatively stable) and remote from the locations of large former ice sheets. The resulting record indicates a large sea-level jump shortly before 13,800 calendar years BP, which corresponds to meltwater pulse 1A in the Barbados coral records^{8,9}. The timing of this event is more accurately constrained in the Tahiti record, revealing that the meltwater pulse coincides with a short and intense climate cooling event¹²⁻¹⁵ that followed the initiation of the Bölling-Allerød warm period¹²⁻¹⁶, but preceded the Younger Dryas cold event by about 1,000 years.

The Barbados sea-level record^{8,17,18} suggested that the last deglaciation was characterized by two brief periods of accelerated melting superimposed on a smooth and continuous rise of sea level with no reversals. These so-called MWP-1A and MWP-1B events (~14,000 and 11,300 calendar years BP, cal. yr BP) are thought to correspond to massive inputs of continental ice: for

^{*} Present address: Observatoire Océanologique Européen de Monaco, Av. St Martin, MC-98000 Monaco.

example, $\sim 50\text{--}40\text{ mm yr}^{-1}$ is roughly equivalent to discharge rates of $16,000\text{ km}^3\text{ yr}^{-1}$ for MWP-1A). A third meltwater pulse, smaller than the two other, was identified at $\sim 7,600\text{ cal. yr BP}$ in a compilation of Caribbean corals together with the Barbados curve¹⁹.

Nevertheless, there are still some doubts concerning the reality of the meltwater pulses (see, for example, ref. 20). Indeed, the Barbados sea-level curve was obtained from three separate cores and within each core a nearly constant sea level rise is observed. Each of these segments is offset from the next, and these offsets constitute the two rapid meltwater surges. A similar problem arises for the third meltwater pulse¹⁹ which corresponds to the junction between the data sets obtained for the Barbados drills and for other Caribbean corals collected by diving. In addition the sea-level jump evidenced in New Guinea ($\sim 11,000\text{ cal. yr BP}$) is delayed by a few centuries when compared¹¹ to the one observed at Barbados. Furthermore, these two islands are located in active subduction zones where vertical movements result from discontinuous motion. Consequently, the assumption of constant uplift rates may only be a first-order approximation: 0.34 mm yr^{-1} for Barbados^{8,9,18} and 1.9 mm yr^{-1} for the Huon peninsula in New Guinea^{10,11}.

Reef cores recently taken by the French ORSTOM team at Tahiti provided an opportunity to obtain a new sea-level record in a completely different tectonic and hydro-isostatic setting than Barbados and New Guinea. Long vertical cores (drill holes P6 and P7, separated by 700 m) taken from the outer barrier reef flat protecting Papeete harbour show a sequence of reef carbonates overlying the basaltic substratum at 114 m core depth (that is, below the present sea level, Fig. 1). The P7 carbonate sequence is divided into two subunits separated by a clear unconformity with evidence of subaerial exposure at $\sim 86\text{--}87\text{ m}$ depth.

$^{230}\text{Th}/^{234}\text{U}$ ages (Fig. 1) were measured with a VG-54-30 mass spectrometer fitted with an ion-counting Daly detector (CEREGE/Aix-en-Provence). The chemical and mass spectrometric techniques are similar to those previously described^{21,22}. The 2σ precision of the $^{230}\text{Th}/^{234}\text{U}$ ages ranges from 30 to 60 years for ages between 8,000 and 14,000 cal. yr BP ($^{230}\text{Th}/^{234}\text{U}$ ages are considered as equivalent to calendar ages¹⁷). The precision of the ages has been checked by measuring numerous replicates (Fig. 1; also see Supplementary Information). In particular, we performed five analyses of different pieces of the same coral specimen (sample P7-7 at $\sim 56\text{ m}$ depth). The five $^{230}\text{Th}/^{234}\text{U}$ ages agree with each other within the 2σ errors, with an overall error on the mean (11,005 cal. yr BP) of 12 yr and a maximum difference between replicates of $\sim 30\text{ yr}$.

Values of $\delta^{234}\text{U}_i$ (that is, $(\text{initial } ^{234}\text{U}/^{238}\text{U} - 1) \times 1,000$) obtained on the Tahiti corals are between 142 and 150‰ with a mean value of 147‰ and a standard deviation of 1.5‰ on 40 measure-

ments. This average is indistinguishable from the values measured on present-day sea water (144‰, standard deviation 4‰)²³ and recent corals^{11,17}. This further confirms that the samples remained closed systems for U–Th in the past. The magnitude of the ^{14}C – ^{230}Th age difference observed on the Tahiti corals agrees well with previous studies of corals^{11,17,24} and varved sequences²⁵.

The species assemblages of several types of organisms (for example, corals, encrusting algae, sessile foraminifers and verme-

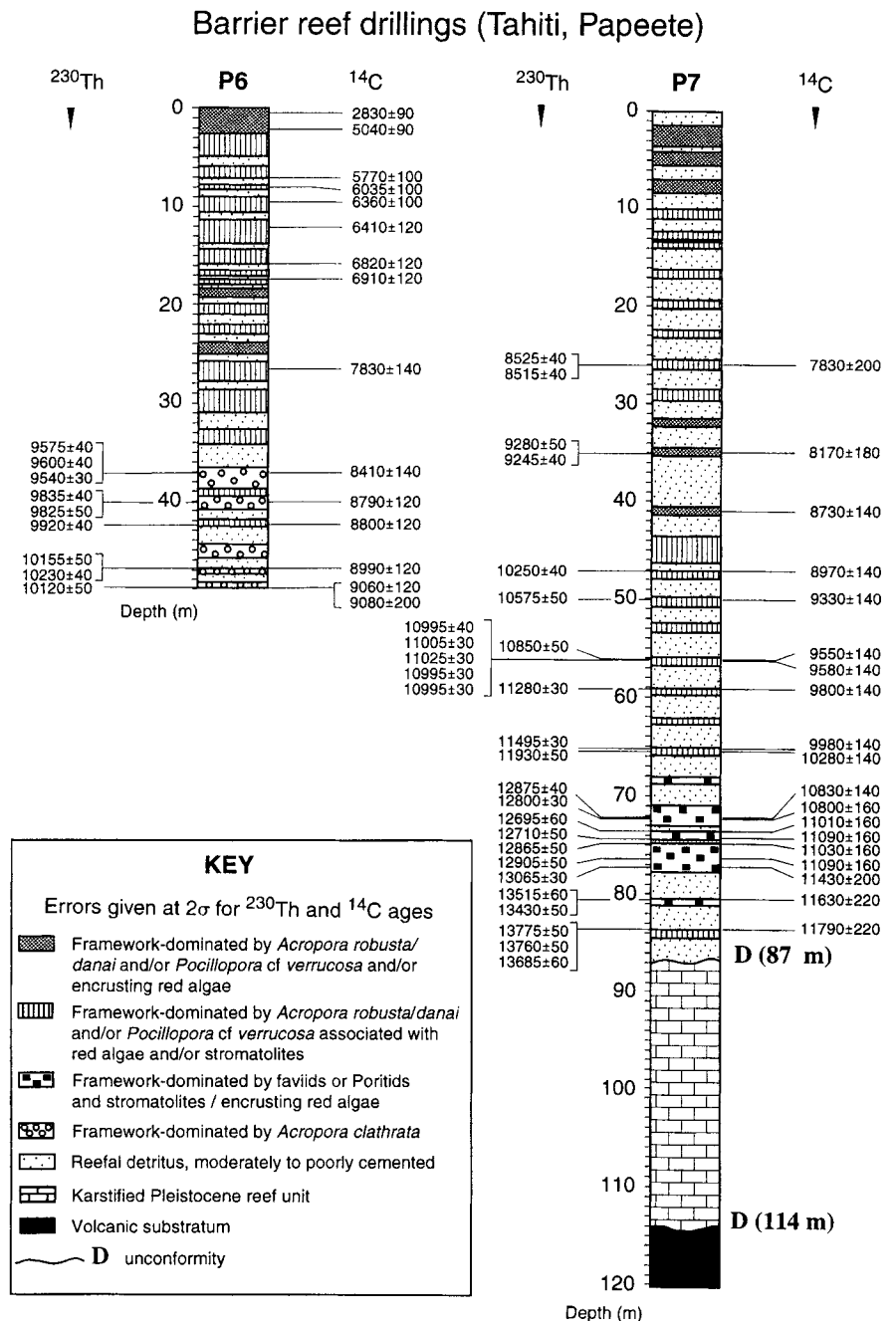


FIG. 1 Simplified representation of cores P6 and P7 from Papeete outer barrier reef flat. Details of the reef sedimentology and palaeoecology will be published elsewhere (L.M. et al., manuscript in preparation). Below the unconformity ('D' at 86–87 m), the Pleistocene subunit exhibits freshwater-induced diagenetic features: karstified reef and the presence of 100% low magnesian calcite rock. In the upper subunit, corals are well preserved and devoid of secondary high magnesian calcite infilling. U–Th and conventional ^{14}C ages are indicated on the left and right sides, respectively. (The ^{14}C ages were measured on the Tandemtron accelerator installed at Gif-Sur-Yvette). Errors are given at 2σ . The replicates are indicated by vertical lines. U–Th ages have been calculated with the following ^{234}U and ^{230}Th decay constants: $\lambda_{234} = 2.835 \times 10^{-6}\text{ yr}^{-1}$ and $\lambda_{230} = 9.195 \times 10^{-6}\text{ yr}^{-1}$. 2σ uncertainties on these values are about 2‰ and 8‰, respectively³⁰. The tabulated data are available as Supplementary Information.

tid gastropods) have been studied in detail on P7 and P6 cores and will be presented elsewhere (L.M. *et al.*, manuscript in preparation). By analogy to the modern biotic reef zonation and ecology, these assemblages are diagnostic of a high water-energy reef front or upper fore-reef zone, at depths less than 6 m below mean sea level.

Figure 2 shows the relative sea-level history derived from the new Tahiti U–Th data and from the other published studies on Barbados^{8,9,18} and New Guinea^{10,11}. The small but systematic shifts can be attributed to different hydro-isostatic responses for the three sites^{2,3}. The Tahiti sea-level curve can also be used to study in detail the timing of the last deglaciation and in particular to assess the existence of sea level jumps.

There is no detectable meltwater pulse at $\sim 7,600$ cal. yr BP, but all three curves exhibit a change of deglaciation rate between 11,500 and 11,000 cal. yr BP, that is, at the time of MWP-1B. The age of the small sea-level step at Tahiti corresponds to the one observed at Barbados^{9,18} but is significantly older by ~ 400 yr than the one evidenced in New Guinea¹¹. However, a straight line fits the Tahiti sea-level curve equally well when taking into account the uncertainty on the palaeodepths of sea-level indicators. This suggests that MWP-1B, if real, was smaller than previously thought^{8,18}.

The last sample above the hiatus of the Tahiti record is dated at $13,750 \pm 30$ cal. yr BP and $11,790 \pm 220$ ^{14}C yr BP. New cores drilled in the Tahiti barrier reef confirm the presence of this discontinuity²⁶. An *Acropora* sample found just above the hiatus in a new core gives an age of $13,850 \pm 40$ cal. yr BP which further supports the results obtained in P7 (the contact with the karstified Pleistocene reef and the oldest coral are both deeper by ~ 7 m than in core P7). These ages correspond precisely to the end of the first jump observed on the Barbados sea-level curve (MWP-1A). Indeed, a similar hiatus was also found in core RGF-12 from the Barbados collection, where a coral dated at $30,230 \pm 160$ cal. yr BP was found just below a subaerial exposure overlaid by a continuous sequence starting at $13,730 \pm 100$ cal. yr BP and $11,780 \pm 180$ ^{14}C yr BP (refs 8, 9, 17, 18). The synchronism of the hiatuses in the Barbados and Tahiti cores provides strong support

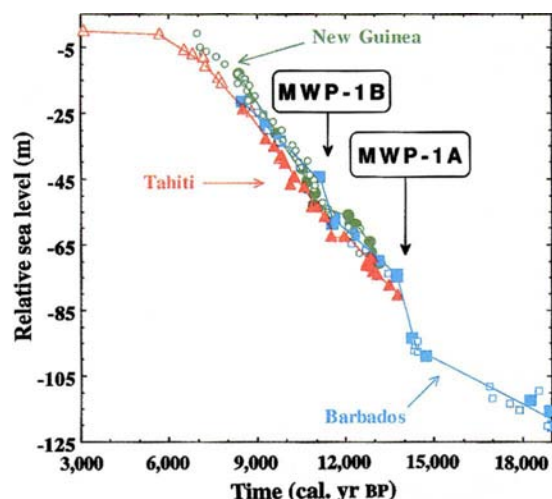


FIG. 2 Sea-level history reconstructed for long drill cores from Tahiti (red triangles), Barbados (blue squares)^{8,9,17,18} and New-Guinea (green dots)^{10,11}. For clarity, samples dated by ^{14}C only^{10,18} have been converted to calendar years by means of calibration formulas²⁴ and are represented with open symbols. Local movements were corrected for the different sites: tectonic uplift for Barbados (~ 0.34 mm yr⁻¹; refs 8, 18) and New Guinea (~ 1.9 mm yr⁻¹; refs 10, 11) and subsidence for Tahiti (~ 0.2 mm yr⁻¹; refs 31, 32). The basalt found below the P7 carbonate sequence has been dated at 549 ± 11 kyr (2σ) by K–Ar (H. Guillou, personal communication). The K–Ar age of the sample belonging to an aerial lava flow is compatible with a subsidence rate of ~ 0.25 mm yr⁻¹ similar to what has been obtained by other authors^{31,32}. Arrows indicate the median ages of MWP-1A and MWP-1B, first evidenced for Barbados^{8,9}.

for the existence and importance of MWP-1A. At both locations older sections ($>14,000$ cal. yr BP) of the sea-level records have to be looked for in cores further offshore. The jump is thus not an artefact due to using different cores to reconstruct sea level²⁰. We show below that the core break at about 14,000 cal. yr BP should be merely considered as an additional argument for a major sea level jump.

We used a numerical model simulating the reef building in order to study the effect of a sea-level jump on the reef geometry. This simple modelling allows qualitative assessment of the effect of sea-level fluctuations on the reef shape and age–depth relationships. For the sake of simplicity, we used a constant slope for the initial topography and a simplified sea-level curve including a major sea-level jump at 14,000 cal. yr BP. Age–depth records in drill holes are calculated by the model. Deglaciation records provided by a series of cores have been simulated for different

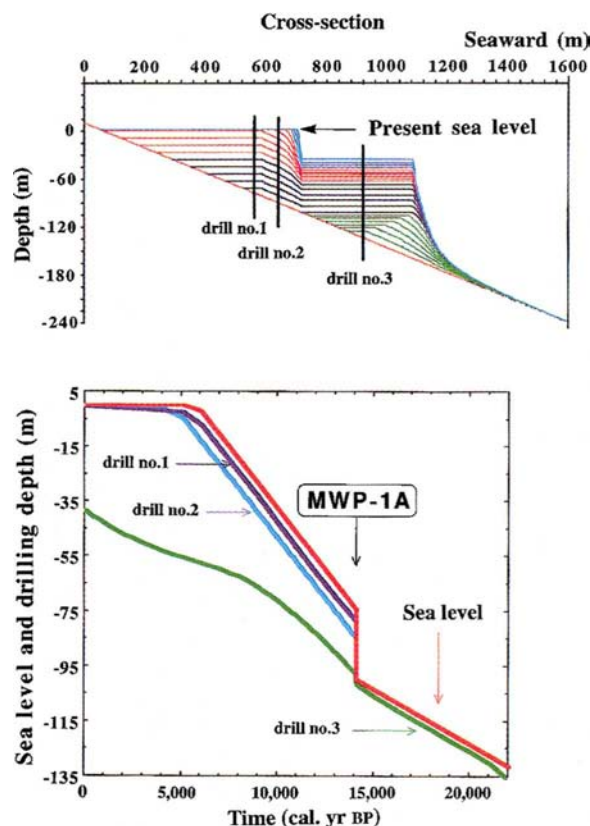


FIG. 3 Top panel, two-dimensional cross-section of a reef accreted on a simple topography (red line ramp) during the past 22,000 cal. yr BP. The x-axis is the distance seaward, and the y-axis is the depth (for graphical clarity the two scales are different). The input sea-level curve is shown on the lower panel (thick red curve). Each curve within the cross-section represents an isochron topography (the time step being 1,000 yr). For clarity, there is a colour code to distinguish the oldest from the youngest corals (green, 22,000–17,000; dark blue, 16,000–11,000; pink, 10,000–5,000; light blue, 4,000 cal. yr BP to the present). Equations of the model are similar to those in ref. 33. Details of the simulations will be presented elsewhere (E.B. *et al.*, manuscript in preparation). Vertical black lines show the positions of three vertical drills simulated by the model (drill no. 1 to drill no. 3). The coral age versus depth relationships in the drill holes are shown on the lower panel (drill no. 1, dark blue; drill no. 2, light blue; drill no. 3, green). Note that the palaeodepth of corals increases when the location of the drill is progressing seawards (drill nos. 1 and 2) but that the hiatus will always be at 14,000 cal. yr BP. Preliminary data on a new Tahiti drill deviated seaward²⁶ seem to confirm this scenario: the contacts with the karstified Pleistocene reef and the oldest coral ($\sim 13,800$ yr BP) are both deeper (~ 7 m) than in core P7. In addition, preliminary data on biological assemblages suggest an increase of palaeodepth, as in the model. In simulation, drill no. 3 recovers deeper-living corals in the reef section corresponding to the past 14,000 cal. yr BP. In reality, these deeper living corals do not form continuous blankets and the chance of recovery is weak.

positions progressing towards the shore (Fig. 3). Up to a threshold, the age–depth relationship of a core always stops, with a hiatus, at a maximum age which corresponds precisely to the major sea-level step (drill holes 1 and 2). Only off-shore drilling enables to recover corals deposited before 14,000 cal. yr BP (drill hole 3). This simple approach clearly illustrates that a major sea-level jump creates a strong unconformity in the reef structure which is in turn responsible for the presence of a synchronous hiatus in cores drilled at different positions on the same reef. The presence of such an unconformity is thus an adequate explanation for the exact synchronism (with both ^{230}Th and ^{14}C chronometers) of the hiatuses observed for Barbados and Tahiti which cannot be accidental. The final step of this project should be to organize an expedition to drill off-shore Tahiti in order to sample corals which were living during the period between the Last Glacial Maximum and MWP-1A.

The existence of a major meltwater pulse shortly before 13,800 cal. yr BP must have dramatically changed the sea surface salinity and temperature distribution of the Atlantic Ocean and hence the thermohaline circulation which maintains a rather mild climate on the European continent. To evaluate the role of the melting history on the North Atlantic climate it is enlightening to plot the meltwater pulses on the same timescale as the Greenland ice $\delta^{18}\text{O}$ profiles¹² and the $\delta^{18}\text{O}$ record of deep-sea core SU81-18²⁷. Figure 4 demonstrates that the sharp temperature rise at ~14,700 cal. yr BP cannot be a consequence of MWP-1A at ~14,000 cal. yr BP, as proposed recently¹⁹ but criticized by Clark²⁸. In fact, MWP-1A probably corresponds to the first major cooling event of the Bølling–Allerød period evidenced at ~13,800–13,900 cal. yr BP in the GRIP core. This event may be equivalent to the Older Dryas climate period²⁹. The same cooling is also very clear at 14,000–14,100 cal. yr BP in the GISPII time series of $\delta^{18}\text{O}$ ¹⁵, ice accumulation rate¹³ and electrical conductivity¹⁴.

The effect of an abrupt meltwater discharge on the temperature of the North Atlantic was recently studied with a coupled ocean–atmosphere general circulation model⁵. During the modelled meltwater pulse the thermohaline circulation and the North

Atlantic surface temperatures weaken abruptly, then exhibit a complex transient behaviour consisting of an abrupt increase and decrease followed by a gradual recovery to the initial state. Similar transient oscillations can be observed (Fig. 4) in the $\delta^{18}\text{O}$ profiles obtained for the Bølling–Allerød/Younger Dryas section of the GRIP ice core¹². Although MWP-1A occurred about a millenium earlier than the start of the Younger Dryas cold period (~12,700 and ~12,900 cal. yr BP in the GRIP and GISPII chronologies), it could have acted as the initial trigger of transient perturbations of the thermohaline circulation of the North Atlantic: decadal–centennial oscillations ultimately leading to a millenium-long cold period, that is, the Younger Dryas event. □

Received 8 January; accepted 29 May 1996.

- Lindstrom, D. R. & MacAyeal, D. R. *Nature* **365**, 214–215 (1993).
- Lambeck, K. *Tectonophysics* **233**, 15–37 (1993).
- Peltier, W. R. *Science* **265**, 195–201 (1994).
- Stocker, T. F. & Wright, D. G. *Nature* **351**, 729–732 (1991).
- Manabe, S. & Stouffer, R. J. *Nature* **378**, 165–167 (1995).
- Rahmstorf, S. *Nature* **378**, 145–149 (1995).
- Schiller, A., Mikolajewicz, U. & Voss, R. *Max-Planck für Meteorology Rep.* **188**, 1–42 (1996).
- Fairbanks, R. G. *Nature* **342**, 637–647 (1989).
- Bard, E., Hamelin, B. & Fairbanks, R. G. *Nature* **346**, 456–458 (1990).
- Chappell, J. & Polach, H. *Nature* **349**, 147–149 (1991).
- Edwards, R. L. et al. *Science* **260**, 962–968 (1993).
- Johnsen, S. J. et al. *Nature* **359**, 311–313 (1992).
- Alley, R. B. et al. *Nature* **362**, 527–529 (1993).
- Taylor, K. C. et al. *Nature* **361**, 432–436 (1993).
- Groote, P. M., Stuiver, M., White, J. W. C., Johnsen, S. & Jouzel, J. *Nature* **366**, 552–554 (1993).
- Broecker, W. S. *Quat. Res.* **38**, 135–138 (1992).
- Bard, E., Hamelin, B., Fairbanks, R. G. & Zindler, A. *Nature* **345**, 405–410 (1990).
- Fairbanks, R. G. *Paleoceanography* **5**, 937–948 (1990).
- Blanchon, P. & Shaw, J. *Geology* **23**, 4–8 (1995).
- Broecker, W. S. *Paleoceanography* **5**, 459–467 (1990).
- Bard, E. et al. *Nucl. Instrum. Meth.* **B52**, 461–468 (1990).
- Bard, E. et al. *Geophys. Res. Lett.* **23**(12), 1473–1476 (1996).
- Chen, J. H., Edwards, R. L. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **80**, 241–251 (1986).
- Bard, E., Arnold, M., Fairbanks, R. G. & Hamelin, B. *Radiocarbon* **35**, 191–199 (1993).
- Goslar, T. M. et al. *Nature* **377**, 414–417 (1995).
- Cabioc, G., Join, Y., Ihilly, C. & Laurent, J.-L. *Rapports de Missions ORSTOM* 32 (ORSTOM, Nouméa, 1995).
- Bard, E. et al. *Nature* **328**, 791–794 (1987).
- Clark, P. U. *Geology* **23**, 957–959 (1995).
- Mangerud, J., Andersen, S. T., Berglund, B. E. & Donner, J. J. *Boreas* **3**, 109–128 (1974).
- Ludwig, K. R. et al. *Science* **258**, 284–287 (1992).
- Pirazzoli, P. A. & Montagnoni, L. F. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **68**, 153–175 (1988).
- Le Roy, I. thesis, Univ. Paris XI (1994).
- Bosscher, H. & Schlager, W. *Sedimentology* **39**, 503–512 (1992).

ACKNOWLEDGEMENTS. We thank P. Clark, M. Field, H. Guillou, J. Guiot, S. Johnsen, J. Jouzel, M. Maia, D. Peteet and J. Récy for fruitful discussions and/or for critical review of the manuscript; B. Angeletti, W. Barthelmy, K. Clavier, P. Dejean, E. Kaltnecker, B. Laude, A. Laurenti, J. J. Motte, N. Tisnerat and K. Van de Velde for technical help; S. Johnsen for numerical data of $\delta^{18}\text{O}$ in the GRIP ice core; H. Guillou for communicating his unpublished K/Ar age on the basalt in core P7. This study was supported by ORSTOM (Endo-Upwelling Program) and PNRC. CEA and CNRS supported the AMS measurements.

SUPPLEMENTARY INFORMATION. Available on Nature's World-Wide Web site <http://www.nature.com>. Paper copies are available from Mary Sheehan at the London editorial office of Nature.

CORRESPONDENCE should be addressed to E.B. (e-mail: ebard@arbois.cerege.fr).

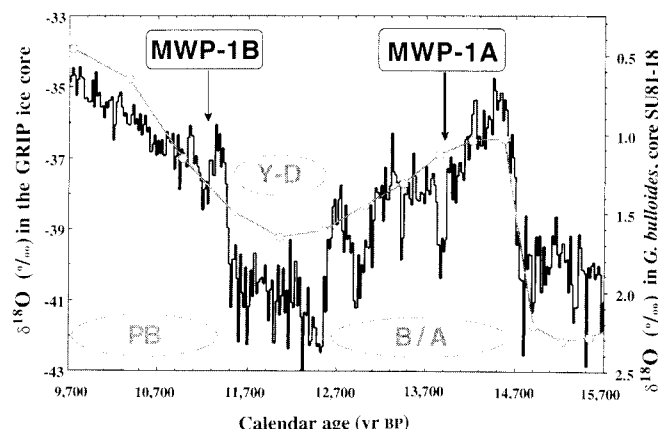


FIG. 4 The solid line represents the $\delta^{18}\text{O}$ variations measured in the GRIP ice core¹² (numerical data provided by S. Johnsen, updated chronology from C. Hammer, personal communication) and the open dots show the planktonic foraminifera $\delta^{18}\text{O}$ record of deep sea core SU81-18 raised at about 38°N in the Northeast Atlantic²⁷. The SU81-18 chronology is based on more than 22 AMS- ^{14}C ages²⁷ converted to calendar years by means of calibration formulas²⁴. The strong warming ($\delta^{18}\text{O}$ increase at ~14,700 cal. yr BP in the Greenland ice) at the beginning of the Bølling–Allerød (B/A) period is synchronous with the sharp warming in core SU81-18 ($\delta^{18}\text{O}$ decrease in the planktonic foraminifera which cannot be due to MWP-1A as proposed previously⁸). MWP-1A dated at ~14,000 cal. yr BP is significantly younger than the strong warming at 14,700 cal. yr BP. It corresponds to the first cold reversal of the Bølling–Allerød observed in the GRIP core (13,800–13,900 cal. yr BP¹²) and GISPII core (14,000–14,100 cal. yr BP¹⁵). MWP-1B (if real) could correspond to a small and abrupt cooling at about 11,300 cal. yr BP in the GRIP ice core¹².

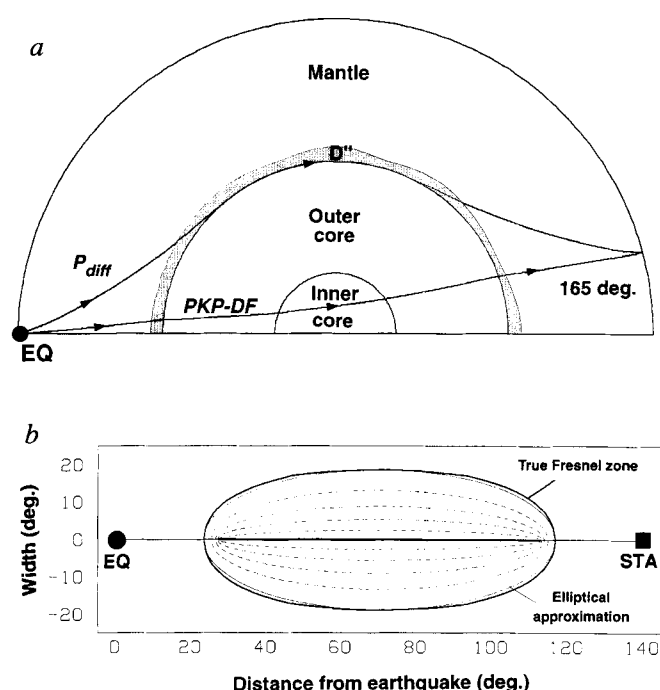
Large-scale structure at the core–mantle boundary from diffracted waves

Michael E. Wysession

Department of Earth and Planetary Sciences, Washington University, St Louis, Missouri 63130, USA

At the base of the Earth's mantle is a region, called D'' , which serves as a thermal and chemical boundary layer between the silicate mantle and the liquid-iron outer core^{1,2}. Tomographic models of mantle compressional-wave velocity (v_p) have their worst resolution in D'' , owing to limited ray sampling^{3–5}, but have hinted at large-scale lateral variations there. Here I use a new technique, which has its greatest resolution within D'' , to produce

FIG. 1 *a*, The geometry of the P_{diff} and PKP-DF seismic phases used to determine the PKP – P_{diff} differential travel times. A total of 543 digital seismograms for earthquake–station distances ranges of 120° to 165° were used from 1980–1987. For an earthquake–station distance of 165° , P_{diff} spends 503 s (45% of its total travel time) within D'' . In contrast, PKP-DF spends 45 s in a 300-km-thick D'' , only 9% of the P_{diff} time. The paths of the P_{diff} waves outside D'' , as well as the complete mantle paths of PKP-DF, are corrected for mantle anomalies by ray-tracing through a whole-mantle tomographic model⁵. Only one-dimensional ray tracing is used, with the ray-path perturbations taken to be second-order contributions. The inner-core paths of PKP-DF are also corrected for azimuthal anisotropy²⁷. Spurious source and receiver effects are removed through the use of differential arrival times for identical earthquake–station pairs, which are obtained through comparison with synthetically generated PKP and P_{diff} phases. *b*, Representation of the 20-s Fresnel zone of a P_{diff} wave at the CMB. The size of the Fresnel zone limits the resolution of the study to large-scale features. The true Fresnel zone (heavy solid line) is computed as the set of reflectors between two minimum time paths (earthquake–reflector and reflector–seismometer) that yields a total time that is a quarter-wavelength (5 s) later than the direct P_{diff} path. This CMB-sampling P_{diff} 'footprint' is computed in three dimensions, but shown schematically in two dimensions. Also shown is an elliptical approximation (thin solid line) and the exponential contours (dashed lines) used for weighting the distribution of the PKP – P_{diff} anomalies. The PKP – P_{diff} residuals are converted into velocity anomalies as $\% \Delta v = 100 \times (t_{\text{tot}} - t_{\text{res}}) / t_{\text{tot}}$ where $\% \Delta v$ is the inferred P-velocity perturbation along the D'' path of P_{diff} and t_{tot} is the travel time of P_{diff} in D'' . Because the long-period P_{diff} Fresnel zones are so large (by 165° they are 7,200 km long and 3,800 km wide), we do not invert the infinite-frequency ray-paths but instead the weighted Fresnel zone distributions. The $\% \Delta v$ anomalies are distributed across the elliptical approximations of the 20-s P_{diff} Fresnel zones, and weighted by distance from their geometrical ray paths according to weight $= (e^{(b-b')/b} - 1) / (e - 1)$ where b is the semi-minor axis of the 20-s Fresnel zone and b' is the semi-minor axis of the dashed contours. The exponential weighting provides an adequate representation of the Fresnel-zone sampling based upon Kirchhoff synthetics (T. Lay, personal commu-



nication), and though the peak energy of P_{diff} is much larger than 20 s at the longer distances, I use a consistent 20-s Fresnel zone to be conservative. These weighted Fresnel zones are superimposed upon a grid of 660 nodes that cover the CMB in a manner that maximizes the minimum distance between nodes (N. Sloane, personal communication).

a map of the large-scale v_p variations within D'' . The technique compares the arrival times of waves that have been refracted across and diffracted around the core–mantle boundary (CMB). The diffracted waves travel a long way in D'' , making them excellent probes of this layer, and the differential technique removes many effects arising from ray paths outside of D'' (ref. 6). The new map provides an image of the continent-sized variations at the CMB that is complementary to existing shear-wave data^{7,8}, and similarities between the map and projections of ancient subducted lithosphere⁹ verify previous suggestions¹⁰ of a strong coupling between surface plate tectonics and the base of the mantle.

The modelling of differential seismic phases has revealed a high level of complexity in D'' , including discontinuous increases in velocity at the top of the region^{2,11–14}, a very-low-velocity layer just above the core–mantle boundary (CMB)¹⁵, and the presence of seismic anisotropy^{16,17}. The structure of D'' may be analogous to the crust and lithosphere at the Earth's surface, and be the result of thermal, chemical and mineralogical phase variations. However, the large-scale P-velocity variations at the CMB are still uncertain: existing tomographic P-velocity models correlate poorly with each other due to a 'severe decrease in resolution' at the base of the mantle⁴, and studies of profiles of diffracted P-waves (P_{diff}) have limited regional coverage^{18,19}. The use of P_{diff} in a differential phase study (Fig. 1*a*) does not require a simultaneous inversion for the rest of the mantle, and gives a large-scale map that is well resolved over most of the globe. Although whole-mantle tomographic models have the potential to map D'' at smaller scales using P, PcP and PKP travel times, adequate resolution is currently limited to the CMB beneath eastern Asia and parts of the Pacific⁴. Shear-velocity models have incorporated ScS and SKS data, and have attained a better resolution of the base of the mantle^{20–24}. However, a full understanding of the geophysical processes operating in D'' will require knowledge of both P and S velocities, and the differences between the two have significance as to the thermochemistry of the deep Earth. For

instance, a decrease in the ratio of S-to-P velocities could mean an increase in the silicate/oxide ratio of D'' (ref. 18).

Diffracted P-waves have not been included in tomographic studies because of the rapid decay of amplitude with increasing frequency during core-diffraction. P_{diff} is infrequently seen on short-period seismograms and therefore rarely reported to agencies such as the International Seismological Centre. The lack of high frequencies for all but the shortest diffraction distances makes picking arrival times unreliable, and the resulting long-period waveform is too dissimilar from other phases in the seismogram to allow for a cross-correlated differential travel time. The present technique uses a comparison between the waveforms of P_{diff} and the refracted phase, PKP-DF, and their synthetic counterparts⁶, computed by using a reflectivity method²⁵ (Fig. 2). The synthetic seismograms are computed with the PREM velocity structure²⁶, and deviations in the timing of the P_{diff} and PKP-DF arrivals represent anomalies from this reference Earth model. The first parts of the pulses are aligned²⁰, to avoid

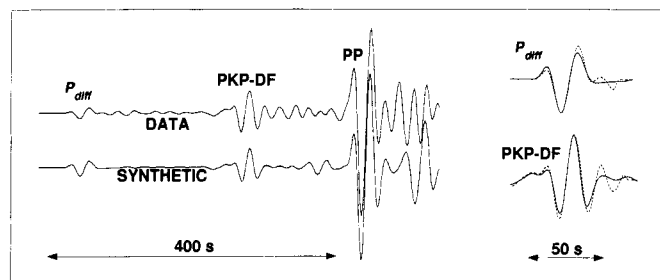


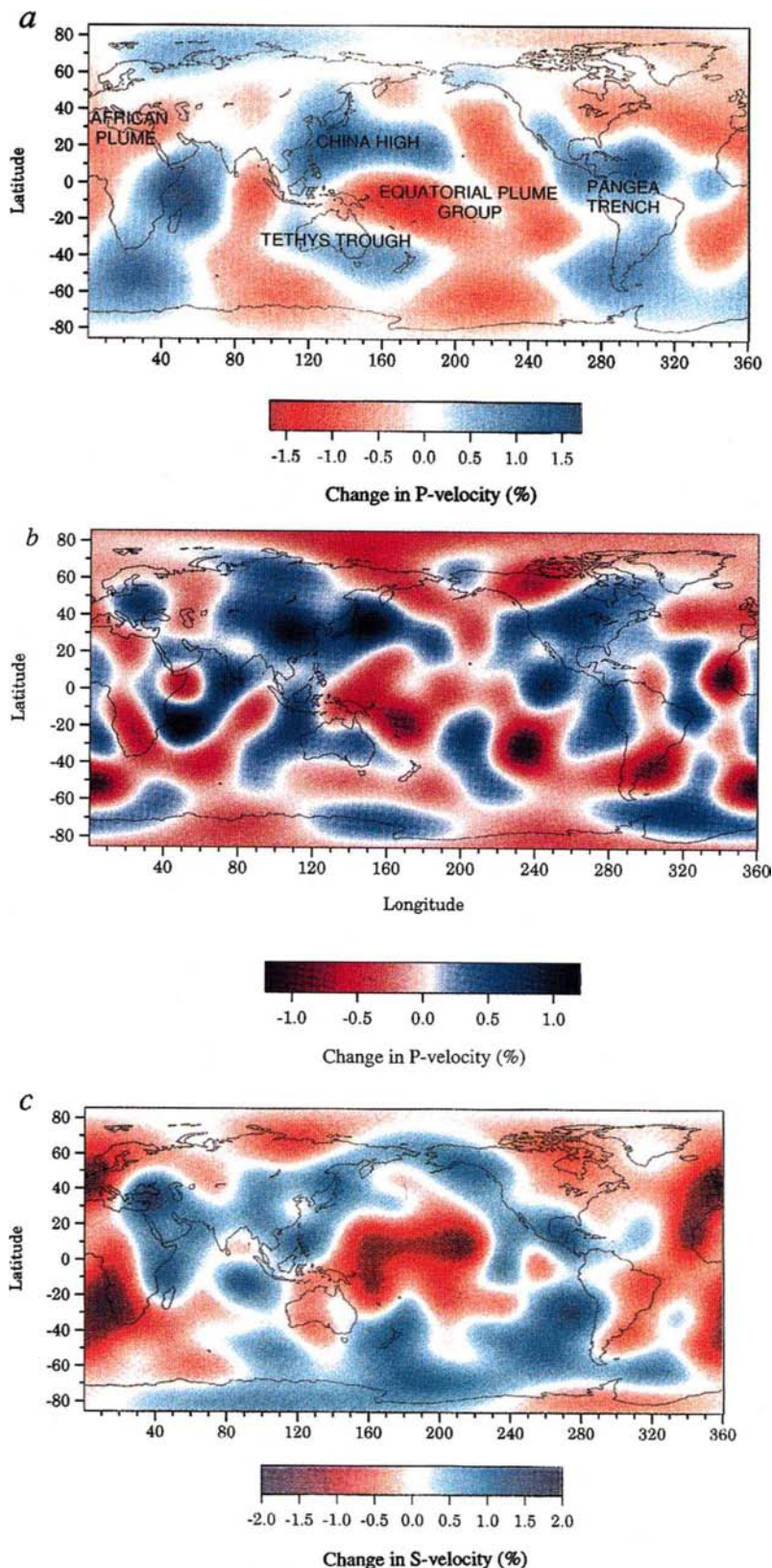
FIG. 2 An example of the data and synthetic seismograms, shown for a Vanuatu earthquake on 5 August, 1982, recorded at station KONO at a distance of 129.8° . At right are shown the separate overlays for the P_{diff} and PKP-DF phases, with the data shown as dashed lines and the synthetics shown as solid lines.

contamination from long source–time functions, and the $PKP - P_{diff}$ travel-time residuals are determined as

$$\begin{aligned}
 tt_{res} &= (PKP - P_{diff})_{residual} \\
 &= (PKP_{data} - PKP_{synth}) - (P_{diff,data} - P_{diff,synth}) + \text{corrections}
 \end{aligned}$$

where travel-time ‘corrections’ are applied to help remove effects from ellipticity, inner-core anisotropy²⁷, the PKP mantle paths, and the P_{diff} mantle paths outside D'' using ray-tracing through a three-dimensional tomographic whole-mantle model⁵. The 543 residuals had a mean of -0.55 s, with a standard deviation of ± 1.87 s and a total range from -6.1 to $+4.1$ s. Variations in CMB topography cause changes to $PKP - P_{diff}$ of

FIG. 3 a, Map of the long-wavelength variation of P-velocities at the base of the mantle. The resolution of features is of the order of $>2,000$ km, based on the widths of the P_{diff} CMB Fresnel zones and synthetic retrieval tests (Fig. 4). The velocities show a 3% lateral variation averaged over the bottom several hundred kilometres of the mantle. The exact vertical sampling range of P_{diff} up into the mantle along its CMB path varies as a function of frequency and CMB distance travelled. The labelled features come from the nomenclature of Dziewonski and colleagues³². There is a striking correlation between regions of fast D'' velocities and the locations of Mesozoic palaeo-trenches⁹. Regions of fast P-velocity at the base of the mantle correspond to surface locations of the subduction of the Panthalassa, Pacific, Farallon and Tethys Sea floors since the Jurassic. Node velocities are obtained using a standard weighted, damped least-squares inversion²⁸ of the $PKP - P_{diff}$ residuals distributed across the CMB Fresnel zones. The damping parameter is taken to be 1.2, which provides a good balance between weighing the least squares for the overdetermined part of the inversion and weighing the minimum length for the undetermined part. The map is formed by interpolating between the node values with a moving gaussian cap. b, Map of D'' P-velocity variations by Vasco and colleagues⁵, using a tomographic inversion of International Seismological Centre (ISC) P-arrival times. The regions beneath eastern Asia and the western Pacific Ocean are the only resolvable part of the lowermost mantle using this approach. Note that these regions correlate well with those shown in a, an encouraging fact given that the two maps are made with independent data. The P-wave data provide almost no coverage of the Southern Hemisphere, which is remedied by the $PKP - P_{diff}$ data. c, Map of D'' S-velocity variations (at a depth of 2,850 km) by Su and colleagues⁷, using a tomographic inversion of waveforms and absolute and differential travel times. This model provided³² the palaeosubduction designations labelled in a. The circum-Pacific ring of fast D'' velocities, surrounding the sub-Pacific slow velocity zone, is well developed and is reproduced at longer wavelengths in the $PKP - P_{diff}$ map of a. Although the agreement between this and a is quite strong, there are some disagreements: the circum-Pacific ring of fast velocities is unbroken beneath the northern Pacific Ocean for S-velocities, but not so for P-velocities; D'' beneath the southern Pacific Ocean is fast for S-velocities but slow for P-velocities; and a region of fast S-velocities beneath the mid-East is found further to the south for P-velocities. Although these differences may be the result of inadequate resolution, they may also indicate thermochemical variations which could affect S- and P-velocity variations differently.



less than 0.02 s because they are both slower when the CMB is elevated. Errors in earthquake depths can add only small errors to tt_{res} : at 140° , a 30-km error in hypocentre depth will change tt_{res} by 0.17 s.

A problem with inverting the residuals for velocity structure along these ray-traces is that the P_{diff} waves have enormous Fresnel zones, or effective sampling regions, at the CMB (Fig. 1b). This is due to the great epicentral distances involved (and resulting geometric spreading) and the large dominant periods of the P_{diff} waveforms (> 20 s). The PKP – P_{diff} residuals are therefore converted into velocity anomalies, distributed across the 20-s P_{diff} Fresnel zones, and then superimposed upon a grid of CMB nodes. The velocity perturbations at the nodes are obtained through a weighted, damped, least-squares inversion²⁸ and shown in Fig. 3.

There is a total lateral variation of $\sim 3\%$ in the D'' P-velocity for the long wavelengths at which we have resolution. This is vertically averaged over the lowermost mantle for the sampling regions of the long-period P_{diff} waves. The wavelengths we are able to resolve

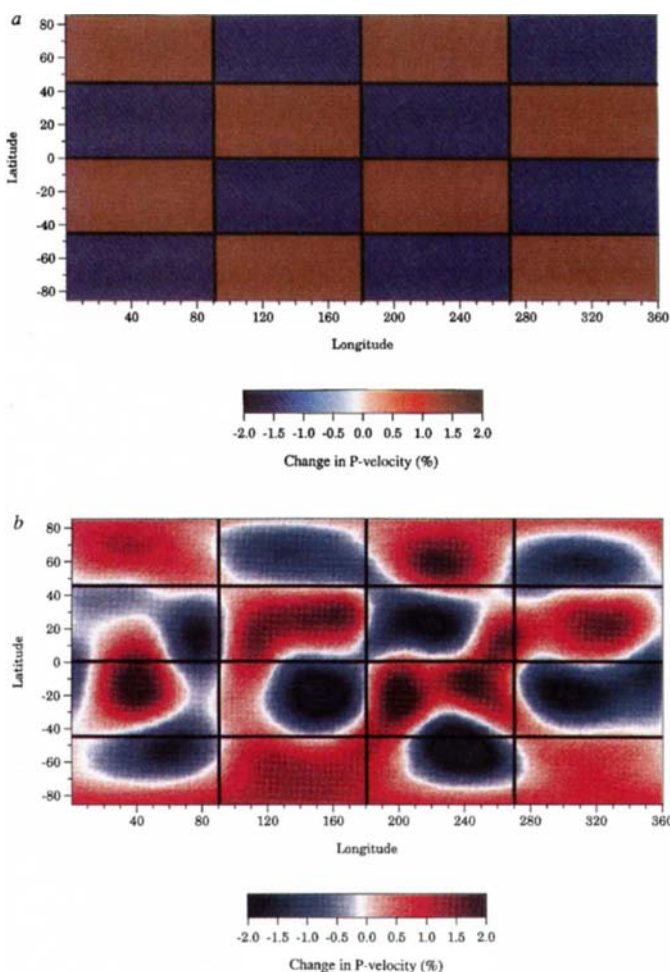


FIG. 4 Recovery of a synthetic model consisting of a chessboard pattern of alternating velocity anomalies. **a**, Each anomaly is originally 45° in latitude by 90° in longitude, and has a value of $\Delta v = \pm 1.5\%$. The one-dimensional ray paths of the 543 data pairs were traced through this chessboard pattern, and the accrued travel-time anomalies were then processed in a manner similar to that of the actual data. **b**, The pattern of recovered anomalies. Most of the CMB has good resolution at the large scales resolved by our P_{diff} paths, although the Northern Hemisphere is better recovered than the Southern Hemisphere. In most cases the approach used here recovers the sharp transitions from fast to slow velocities fairly well. The recovered anomalies had a total range from -2.9% to 2.4% , although more than 90% of the regions were between -2.0 and 2.0% .

are of the order of 2,000 km, based on the *a priori* use of the 20-s P_{diff} Fresnel zones, which have a CMB width on this scale, and the ability to recover a synthetic set of velocity anomalies (Fig. 4). While in general the Northern Hemisphere is better resolved than the Southern Hemisphere, there is good resolution of very large features over most of the globe. The 3% magnitude of variation is comparable with the results of tomographic P-velocity models (2–3%)^{3–5} and differential P_{diff} studies^{18,19}. As the scaling between P and S velocities is likely to be at least 2.5 at the base of the mantle²⁹, the 3% variations are also compatible with a recent long-wavelength tomographic S-velocity model that includes SKS – S_{diff} differential times (analogous to the PKF – P_{diff} here)⁸, but larger than for models that do not include the core-diffracted data^{5,7}.

The striking similarity between the locations of fast seismic velocities in D'' and the locations of palaeosubduction zones⁹ reinforces the idea that D'' may be the graveyard for subducted lithosphere. Much has been said about the possibility of slabs penetrating into the lower mantle³⁰, with some of the strongest support coming from high-resolution tomographic images^{22,31}. Some of the positive seismic anomalies in D'' from tomographic images have previously been given names like the 'Pangea Trench' and 'Tethys Trough' by Dziewonski and colleagues³², suggesting surface origins. Complementary to the idea of fast D'' anomalies corresponding to cold rock brought from the upper mantle is the concept of D'' as the origin of hotspot mantle plumes³³. The PKP – P_{diff} data suggest the locations of two large low-velocity regions that correspond with the Equatorial Pacific plume group and Great African plume features previously identified. The western part of the sub-Pacific low-velocity region has been observed and examined by previous differential phase studies^{20,21,23}, and is one region recently found to be underlain by a narrow zone of even slower velocities¹⁵. The western African plume modelled here extends westwards beneath the northern Atlantic Ocean. I also found a region of very slow velocities beneath the central Indian Ocean.

Figure 3b shows the tomographic image of lowermost-mantle P-velocities of Vasco and colleagues⁵; in the regions of reliable resolution (beneath eastern Asia and the western Pacific Ocean), the pattern correlates very well with the PKP – P_{diff} image in Fig. 3a. It is significant, however, that variations exist between our map and the tomographic image of D'' S-velocities by Su and colleagues⁷, shown in Fig. 3c. Although there is generally very good agreement, such as the indication of the circum-Pacific fast velocities surrounding the sub-Pacific slow region, there are some noticeable differences. It may be that these models still lack the resolution to allow for accurate comparisons, but large variations in Poisson ratio exist in the surface crust as a result of textural and compositional variations³⁴, and it may be the case that variations in the temperature and chemistry of D'' cause variations in the Poisson ratio of these regions¹⁸.

The fact that large-scale continent-sized P-velocity features do exist in D'' , and are not artefacts of inversion techniques, has other significant implications. The thermal or topographic anomalies that cause the velocity variations would be large enough to constrain patterns of outer-core iron flow and therefore our geomagnetic field³⁵. The existence of large continent-sized D'' lateral variations suggests that the scales of convection at the top and bottom of the mantle may be similar, which provides a constraint for mantle convection models. The new determination of D'' P-velocities reported here, with a strong resemblance to the locations of palaeosubduction at the surface⁹, strengthens the existing argument for whole-mantle mass transfer between the surface and CMB. Of course, the similarities could be coincidental or the result of thermal coupling between separate convective layers. If whole-mantle convection occurs episodically³⁶ and we are between episodes, the approach used here might not be able to distinguish between the remnants of subducted slabs, the result of lower-mantle convection, or part of a static chemical boundary layer. □

Received 5 February; accepted 3 June 1996.

1. Loper, D. E. & Lay, T. J. *geophys. Res.* **100**, 6397–6420 (1995).
2. Lay, T. *Eos* **70**, 49, 54–55, 58–59 (1989).
3. Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y. *Phys. Earth planet. Inter.* **59**, 294–328 (1990).
4. Pulliam, R. J., Vasco, D. W. & Johnson, L. R. *J. geophys. Res.* **98**, 699–734 (1993).
5. Vasco, D. W., Johnson, L. R., Pulliam, R. J. & Earle, P. S. *J. geophys. Res.* **99**, 13727–13755 (1994).
6. Wyssession, M. E., Valenzuela, R. W., Bartók, L. & Zhu, A.-N. *Phys. Earth planet. Inter.* **92**, 67–84 (1995).
7. Su, W., Woodward, R. L. & Dziewonski, A. M. *J. geophys. Res.* **99**, 6945–6981 (1994).
8. Liu, X.-F., Su, W.-J. & Dziewonski, A. M. *Eos* (Spring mtg suppl.) **75** (16), 232 (1994).
9. Ricard, Y., Richards, M., Lithgow-Bertelloni, C. & Le Stunff, Y. *J. geophys. Res.* **98**, 21895–21909 (1993).
10. Davies, G. F. *J. geophys. Res.* **89**, 6017–6040 (1984).
11. Revenaugh, J. & Jordan, T. H. *J. geophys. Res.* **96**, 19811–19824 (1991).
12. Howard, S. & Nataf, H.-C. *Phys. Earth planet. Inter.* **72**, 264–275 (1992).
13. Vidale, J. E. & Benz, H. M. *Nature* **361**, 529–532 (1993).
14. Kendall, J. M. & Shearer, P. M. *J. geophys. Res.* **99**, 11575–11590 (1994).
15. Gahmer, E. J. & Helmberger, D. V. *Geophys. Res. Lett.* **23**, 977–980 (1996).
16. Maupin, V. *Phys. Earth planet. Inter.* **87**, 1–32 (1994).
17. Kendall, J.-M. & Silver, P. G. *Nature* **381**, 409–412 (1996).
18. Wyssession, M. E., Okal, E. A. & Bina, C. R. *J. geophys. Res.* **97**, 8749–8764 (1992).
19. Souriau, A. & Poupinet, G. *Phys. Earth planet. Inter.* **84**, 227–234 (1994).
20. Woodward, R. L. & Masters, G. *Nature* **352**, 231–233 (1991).
21. Wyssession, M. E., Bartók, L. & Wilson, J. J. *geophys. Res.* **100**, 8351 (1995).
22. Grand, S. P. *J. geophys. Res.* **99**, 11591–11622 (1994).
23. Silver, P. G. & Bina, C. R. *Geophys. Res. Lett.* **20**, 1135–1138 (1993).
24. Gahmer, E. J. & Helmberger, D. V. *J. geophys. Res.* **98**, 8225–8241 (1993).
25. Kennett, B. L. N. *Seismic Wave Propagation in Stratified Media* (Cambridge Univ. Press, Cambridge, 1983).
26. Dziewonski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297–356 (1981).
27. Song, X. & Helmberger, D. V. *Geophys. Res. Lett.* **20**, 2591–2594 (1993).
28. Menke, W. *Geophysical Data Analysis: Discrete Inverse Theory* 54–55 (Academic, San Diego, 1989).
29. Robertson, G. S. & Woodhouse, J. H. *Eos* (Fall mtg suppl.) **76** (46), 385 (1995).
30. Lay, T. A. *Rev. Earth planet. Sci.* **22**, 33–61 (1994).
31. Van der Hilst, R. D. *Nature* **374**, 154–157 (1995).
32. Dziewonski, A. M., Forte, A. M., Su, W.-J. & Woodward, R. L. in *Relating Geophysical Structures and Processes: The Jeffreys Volume* (eds Aki, K. & Drnowska, R.) 67–105 (Am. Geophys. Union, Washington DC, 1993).
33. Olson, P., Schubert, G. & Anderson, C. *Nature* **327**, 409–413 (1987).
34. Zandt, G. & Ammon, C. J. *Nature* **374**, 152–154 (1995).
35. Bloxham, J. & Jackson, A. *Geophys. Res. Lett.* **17**, 1997–2000 (1990).
36. Tackley, P. J., Stevenson, D. J., Glatzmaier, G. A. & Schubert, G. *Nature* **361**, 699–704 (1993).

ACKNOWLEDGEMENTS. I thank E. Okal and T. Lay for helpful discussions, and C. Bina and T. Clarke for assistance with reflectivity codes. I also thank E. Gahmer for comments and suggestions.

CORRESPONDENCE should be addressed to M.E.W. (e-mail: michael@wucore.wustl.edu).

Soft-bodied fossils from a Silurian volcanoclastic deposit

Derek E. G. Briggs*, David J. Siveter† & Derek J. Siveter‡

* Department of Geology, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, UK

† Department of Geology, University of Leicester, Leicester LE1 7RH, UK

‡ Geological Collections, The University Museum, Parks Road, Oxford OX1 3PW, UK

Fossil deposits that preserve lightly sclerotized and soft-bodied organisms are fundamentally important to our understanding of the history of life on Earth. They provide a much more complete record of ancient communities than does the normal shelly fossil record. Conditions during the Cambrian may have favoured the preservation of soft-bodied organisms¹. Burgess Shale-type^{2–5} and Orsten-type⁶ faunas are becoming increasingly known from this roughly 40-million-year-long period for which we have a growing body of data on the metazoan radiation. Soft-bodied organisms are much less well represented in the subsequent 100 million years. The discovery of a new Silurian soft-bodied biota therefore has the potential to fill an important gap in our knowledge. The relatively deep-water marine environment represented is dominated by previously undiscovered arthropods and polychaetes. Here we describe a group of soft-bodied fossils from

carbonate concretions within a volcanic ash, identifying an important new source of soft-bodied taxa.

Between the Cambrian and the Devonian, the occurrence of soft-bodied fossils is very rare. Only two Ordovician units, the Soom Shale in South Africa⁷ and Beecher's Trilobite Bed in New York state⁸, approach the extent of soft-tissue preservation prevalent in Burgess Shale-type deposits. In the Silurian most deposits preserving non-biomineralized taxa (apart from graptolites) yield only eurypterids and phyllocarid crustaceans, arthropods with thick organic cuticles, and are confined to hypersaline peritidal environments, at least in North America⁹. Soft-bodied organisms preserving mineralized muscle and other labile tissues are known only from the peritidal Llandovery–Wenlock series of Waukesha, Wisconsin^{9,10}, and the restricted shallow marine Llandovery of Lesmahagow in Scotland¹¹. The new Silurian locality yielding soft-bodied fossils, from England, has a fully marine setting.

The new soft-bodied fauna was recovered from 2–20 cm concretions ($n = 378$) that occur in a roughly 1-m-thick volcanic ash-rich deposit in the lower Silurian Wenlock Series (424–430 Myr BP; ref. 12) of Herefordshire. The fauna (Fig. 1) is dominated by a small arthropod (29% of specimens) and polychaete worms (at least 27%), with smaller numbers of trilobites (8%), graptolites (4%), high-spined gastropods (2%) and undetermined arthropod and enigmatic soft-bodied fossils (Fig. 1h). Microfossils include abundant radiolarians (Fig. 1d), the first to be discovered in the Silurian of the UK, and acritarchs. The concretions are dominantly calcitic with some quartz and clay minerals, and accessory iron oxide. Many of the concretions (46%) yield macroinvertebrates, the great majority of which are soft-bodied or lightly sclerotized organisms. The fossiliferous unit lies near the local base of several metres of calcareous shales. These carbonate-rich muds, and the associated volcanic ash, were deposited in relatively deep water west of a shelf area (the Midland Platform) which occupied much of central England and contiguous areas of the Welsh borderland during early to mid-Silurian times^{13,14,15}.

The fossils, including the worms, are three-dimensional and uncompact. As the muscle of polychaetes decays to fluid within a few days¹⁶, resulting in collapse of the carcass, mineralization must have been remarkably rapid. A layer of fine-grained crystalline material about 100 μm thick, essentially identical in composition to the rest of the concretion, occurs around the fossil worms (which are 1–7 mm in diameter). This layer is presumably the result of rapid precipitation of calcite. The rarity of clay minerals within the concretions indicates that calcite precipitated before significant dewatering, while substantial pore space was still available. This notion is supported by their spherical shape. The only concentration of clay minerals occurs as a veneer less than 100 μm thick on parts of the fossils. This may represent sediment, adhering to the skin of the worms, for example. No details of the internal soft-tissues are preserved. The carcasses clearly decayed internally, leaving a void. More than 90% of the infill consists of euhedral calcite crystals. An iron-rich film follows some of the grain boundaries. Rare tiny grains of pyrite occur on some of the crystal surfaces. No phosphorus was detected. The high proportion of concretions that yield soft-bodied fossils indicates that CaCO_3 precipitation was initiated as a result of decay enhanced saturation. The decay of a carcass produced excess HCO_3^- , resulting in the formation of a layer of CaCO_3 around the carcass. HCO_3^- that escaped this surface-controlled precipitation¹⁷ resulted in an extension of CaCO_3 saturation beyond the carcass (the volcanic ash deposit contains 2–5% Ca), promoting continued growth of the concretion¹⁸.

The small arthropod which dominates the fauna so far discovered is 3–4 mm long (Fig. 1a, c, e). With the evidence presently available, we cannot assign this new arthropod to the trilobites or any living chelicerate or crustacean taxon. It represents an intermediate morphology like those that dominate the Burgess Shale. The body is divided into three tagmata. The semicircular

cephalic shield has a small triangular ('cardiac lobe') region at the rear, and a median axial ridge. The head bears a pair of stalked eyes and at least five, probably six, pairs of robust, apparently uniramous head appendages. The thorax consists of three somites, each with a pair of slender biramous appendages. Similar appendages are present at the anterior of the pygidium, which consists of four somites plus a terminal division. The pygidium bears a pair of lateral spines and a median terminal spine. The morphology of the weakly trilobed dorsal exoskeleton is similar to that of nektaspid trilobites like *Soomaspis*¹⁹, but the large number of head appendages, the pedunculate eyes, and pendant limbs are crustacean characters.

The new fauna falls in a gap in the record of polychaetes of about 140 million years, from the Cambrian Burgess Shale of Canada to the Devonian Hunsrück Slate of Germany¹⁶. During this interval polychaetes are known only from scolecodonts (jaw elements), apart from a single specimen from the Ordovician Trenton Limestone of Canada²⁰. The Herefordshire fauna includes at least three new polychaetes (Fig. 1f, g). The first (Fig. 1f) has short widely spaced parapodia, short chaetae, and body annulations with pronounced relief. The posterior section of the body is significantly wider than the anterior. The second (Fig. 1g) is broad with short laterally projecting chaetose parapodia. Chaetae also overlap the body, laterally at least; they may have been attached to a second branch of the parapodia. The third is elongate with simple regularly and closely spaced parapodia with chaetae. Specimens preserve details of the cuticle surface and, in some cases, the gut.

Apart from the soft-bodied fossils, the concretions yield similar shelly and other fossils (with the exception of small brachiopods) to the overlying shales: the calymenid trilobite *Tapinocalymene*²¹, small gastropods, as well as monograptid graptolites and acritarchs. The shales yield the low diversity *Visbyella* community¹⁵ which is characteristic of locations at the outer limits of the shelly benthos during Wenlock time and which has been interpreted as reflecting conditions at depths of about 200 m, at the shelf margin²². The preservation of readily decayed polychaetes indicates that the marine community buried in this volcanoclastic sediment is likely to provide a relatively complete record of the animals present in this deeper water Silurian setting.

Calcite infilled three-dimensional soft-bodied fossils are known from a small number of other localities, notably the Cretaceous Santana formation of Brazil²³. Most three-dimensionally preserved soft-tissues, however, are mineralized in apatite. Other volcanoclastic sequences have yielded soft-bodied fossils, for example, the Miocene lacustrine deposit at Clarkia^{24,25}, Idaho, and the Visean of East Kirkton, Scotland²⁶. These, however, are terrestrial rather than marine in origin and the fossils are not preserved in concretions. Volcanoclastic deposits are widespread in the Silurian of the Anglo-Welsh Basin^{27,28} and elsewhere. The discovery of exceptional preservation in such a deposit therefore reveals an important new source of soft-bodied fossils. □

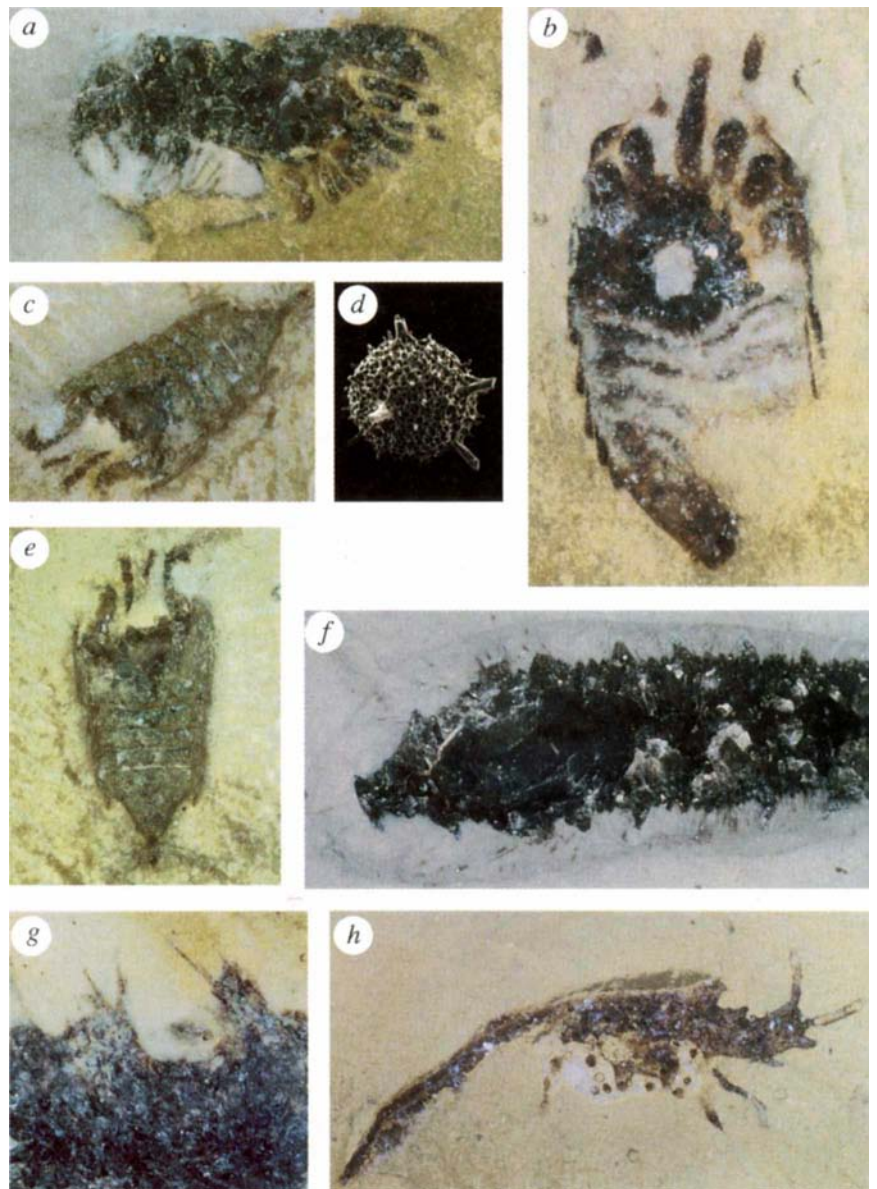


FIG. 1 Fossils from concretions, Silurian (Wenlock Series), Herefordshire. a–c, e, New arthropod. a, Longitudinal perpendicular section, $\times 15.6$ (Oxford University Museum specimen number OUM C.29500a). b, Longitudinal horizontal section, $\times 15.6$ (Leicester University Geology Department, LEIUG 116110). c, e, Anterior oblique and dorsal external views, $\times 8.5$ (OUM C.29501a). d, Astrospherid radiolarian, $\times 23$ (OUM C.29502). f, g, Polychaete worms. f, Type 1, longitudinal horizontal section, $\times 6.5$ (OUM C.29503a). g, Type 2, detail of chaetose parapodia and cuticle, longitudinal horizontal section, $\times 11$ (OUM C.29504a). h, Possible lobopod, $\times 5.5$ (LEIUG 116109).

Received 29 March; accepted 5 June 1996.

1. Allison, P. A. & Briggs, D. E. G. *Geology* **21**, 527–530 (1993).
2. Conway Morris, S. *Palaeontology* **29**, 423–467 (1986).
3. Butterfield, N. J. *Lethaia* **28**, 1–13 (1995).
4. Nedin, C. *Mem. Aus. Ass. Palaeont.* **18**, 31–40 (1995).
5. Hou, X., Ramsköld, L. & Bergström, J. *Zoologica Scripta* **20**, 395–411 (1995).
6. Müller, K. J., Walossek, D. & Zakharov, A. *Neues Jb. Geol. Paläont. Mh.* **197**, 101–118 (1995).
7. Gabbott, S. E., Aldridge, R. J. & Theron, J. N. *Nature* **274**, 800–803 (1995).
8. Briggs, D. E. G., Bottrell, S. H. & Raiswell, R. *Geology* **19**, 1221–1224 (1991).
9. Kluesendorf, J. *Lethaia* **27**, 337–344 (1994).
10. Mikulic, D. G., Briggs, D. E. G. & Kluesendorf, J. *Science* **228**, 715–717 (1985).
11. Walton, E. K. & Oliver, G. J. H. in *Geology of Scotland* (ed. Craig, G. Y.) 195–228 (Geological Society of London, 1991).
12. Harland, W. B. *et al.* *A geological time scale—1989* (Cambridge Univ. Press, 1990).
13. Bassett, M. G. *Palaeontology* **17**, 745–777 (1974).
14. Bassett, M. G., Bluck, B. J., Cave, R., Holland, C. H. & Lawson, J. D. in *Atlas of Palaeogeography and Lithofacies* (eds Cope, J. C. W., Ingham, J. K. & Rawson, P. F.) 37–56 (Geol. Soc. Lond. Mem. **13**, 1992).
15. Hurst, J. M., Hancock, N. J. & McKerrow, W. S. *Proc. Geol. Ass.* **89**, 197–226 (1978).
16. Briggs, D. E. G. & Kear, A. J. *Paleobiology* **19**, 107–135 (1993).

17. Raiswell, R. *Sedimentology* **35**, 571–575 (1988).
18. Raiswell, R. in *Taphonomy, releasing the data locked in the fossil record* (eds Allison, P. A. & Briggs, D. E. G.) 411–453 (Plenum, New York, 1991).
19. Fortey, R. A. & Theron, J. N. *Palaeontology* **37**, 841–861 (1995).
20. Pickerill, R. K. & Forbes, W. H. *Can. J. Earth Sci.* **15**, 659–664 (1978).
21. Siveter, D. J. *Palaeontology* **23**, 783–802 (1980).
22. Brett, C. E., Boucot, A. J. & Jones, B. *Lethaia* **26**, 25–40 (1993).
23. Martill, D. M. *Palaeontology* **31**, 1–18 (1988).
24. Smiley, C. J. & Rember, W. C. in *Communities of the past* (eds Gray, J., Boucot, A. J. & Berry, W. B. N.) 551–590 (Hutchinson Ross, Stroudsburg, 1981).
25. Smiley, C. J. & Rember, W. C. in *Late Cenozoic history of the Pacific Northwest* (ed. Smiley, C. J.) 11–30 (AAAS, San Francisco, 1985).
26. Clarkson, E. N. K., Panchen, A. L. & Rolfe, W. D. I. (eds) *Transactions of the Royal Society of Edinburgh: Earth Sciences* **84**, 175–464 (1994 for 1993).
27. Ross, R. J. *et al. Geol. Mag.* **119**, 135–153 (1982).
28. Teale, C. T. & Spears, D. A. *Sedimentology* **33**, 757–765 (1986).

ACKNOWLEDGEMENTS. We thank R. J. King who found the first fossiliferous nodules and drew the locality to our attention; R. J. Aldridge, S. Conway Morris and W. J. Kennedy for reading an earlier draft of the manuscript; R. G. Clements for discussion and access to material; H. P. Powell for help in the field; A. Smith, A. Swift, S. Kearns, J. B. Cooke and J. Hay for analytical and technical assistance; and P. R. Wilby and M. A. Wills for advice and discussion.

CORRESPONDENCE and requests for materials should be addressed to Derek Siveter. (e-mail: derek.siveter@earth.ox.ac.uk).

Role of leptin in the neuroendocrine response to fasting

Rexford S. Ahima*, Daniel Prabakaran*, Christos Mantzoros*, Daqing Qu†, Bradford Lowell*, Eleftheria Maratos-Flier† & Jeffrey S. Flier*

* Department of Medicine, Division of Endocrinology, Beth Israel Hospital, and † Joslin Diabetes Center, Harvard Medical School, 330 Brookline Avenue, Boston, Massachusetts 02215, USA

A TOTAL deficiency in or resistance to the protein leptin causes severe obesity^{1–4}. As leptin levels rise with increasing adiposity in rodents⁵ and man^{6,7}, it is proposed to act as a negative feedback 'adipostatic signal' to brain centres controlling energy homeostasis, limiting obesity in times of nutritional abundance^{1,3}. Starvation is also a threat to homeostasis that triggers adaptive responses^{8–12}, but whether leptin plays a role in the physiology of starvation is unknown. Leptin concentration falls during starvation¹³ and totally leptin-deficient *ob/ob* mice have neuroendocrine abnormalities similar to those of starvation¹⁴, suggesting that this may be the case. Here we show that preventing the starvation-induced fall in leptin with exogenous leptin substantially blunts the changes in gonadal, adrenal and thyroid axes in male mice, and prevents the starvation-induced delay in ovulation in female mice. In contrast, leptin repletion during this period of starvation has little or no effect on body weight, blood glucose or ketones. We propose that regulation of the neuroendocrine system during starvation could be the main physiological role of leptin.

Depriving male mice of food for 48 h caused a 16% fall in weight (Table 1) and reduced circulating leptin from 3.87 ± 0.35 to 1.5 ± 0.11 ng ml⁻¹ (Fig. 1a). Blood glucose concentration fell from 159 to 63.4 mg dl⁻¹, insulin levels fell from 1.23 to 0.18 ng ml⁻¹ and β -hydroxybutyrate rose from 2.85 to 19.3 mg dl⁻¹ (Table 1). Recombinant leptin injected intraperitoneally (i.p.) into fasted mice produced peak serum concentrations after 1 h, had an apparent half-life of 3 h, and concentrations 6 and 12 h after injection were similar to fed controls (Fig. 1b). When fasted mice were given twice-daily i.p. injections of leptin (1 μ g per g body weight), circulating levels after 48 h were similar to fed controls (Fig. 1a), and leptin repletion did not alter weight loss or levels of glucose, insulin or β -hydroxybutyrate (Table 1). On recommencing feeding, the compensatory rise in food intake was marked during the first 6 h (Table 1). Leptin injection caused a

reduction in food intake over the first 4 h of refeeding, and weight gain after 24 h (Table 1).

Starvation disrupts fertility in humans, non-human primates and rodents^{11,15,16}, most probably at the level of the hypothalamus, as starvation decreases gonadotropin pulse frequency, and levels of gonadotropins and sex steroids¹⁵. We characterized the effect of leptin treatment on the oestrus cycle of mice. Food deprivation for 48 h prolonged dioestrus in all saline-treated mice and delayed vaginal oestrus by 6.3 ± 1.3 days (Fig. 2a), despite 98% of body weight being regained within 24 h of recommencing feeding. In contrast, 7 of 10 leptin-treated fasted mice had no delay in onset of oestrus, and dioestrus was prolonged by 1–3 days in 3 others (Fig. 2a). As in males, leptin did not alter weight loss during starvation, but decreased the rate of weight gain for 24 h after refeeding (data not shown). We found equivalent effects in males, where 48 h of fasting caused an 88% decrease in serum testosterone concentrations from 1.09 ± 0.18 to 0.12 ± 0.02 ng ml⁻¹ (Table 2). Treatment with leptin decreased the fall to 58% of control levels. Serum luteinizing hormone was measurable in 8 of 13 fed mice (1.29 ± 0.40 ng ml⁻¹) and undetectable in the other 5. Luteinizing hormone was undetectable in all 13 fasted mice, and leptin treatment restored levels to nearly 40% of those in fed controls in the 6 of 13 mice where it was detectable (Table 2), indicating that leptin acts at the level of the hypothalamic-pituitary system.

The changes in adrenal and thyroid axes induced by starvation are also influenced by leptin. Serum thyroxine decreased from 0.85 ± 0.05 to 0.18 ± 0.02 μ g dl⁻¹ with fasting, and leptin

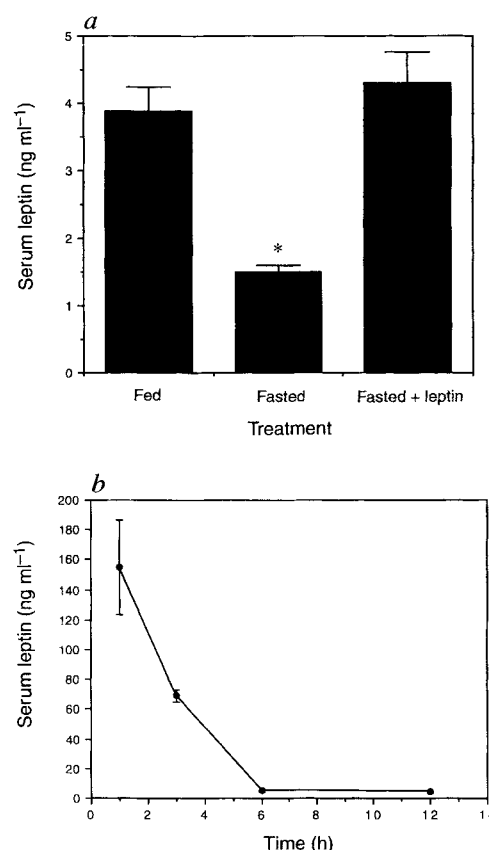


FIG. 1 a, Serum leptin was measured in fed, fasted and fasted + leptin-treated mice (see Table 1 for details) with a radioimmunoassay kit to mouse leptin (Linco Research Inc.). Data are means $\pm$ s.e.m., $n = 12$ per group. * $P < 0.05$ by ANOVA and Fisher PSLD. b, Kinetics of intraperitoneal (IP) recombinant leptin in fasted mice. Serum leptin was measured 1, 3, 6 and 12 h after i.p. injection of recombinant mouse leptin into mice fasted for 48 h; $n = 5$ per group.

TABLE 1 Effects of fasting and leptin treatment

	Fed	Fasted	Fasted + leptin
Body weight (g)			
Initial	24.5 ± 0.5	24.8 ± 0.4	24.7 ± 0.3
After fasting	24.9 ± 0.3	20.7 ± 0.4*	20.3 ± 0.3*
Refeeding 24 h	24.7 ± 0.5	23.8 ± 0.5	22.8 ± 0.3
48 h	24.6 ± 0.7	24.9 ± 0.4	24.8 ± 0.5
Change in weight after 24 h refeeding (g)	0.2 ± 0.1	3.2 ± 0.1*	2.6 ± 0.2*†
Blood glucose (mg dl ⁻¹)	159 ± 5.5	63.4 ± 1.1*	61.1 ± 2.6*
β-hydroxybutyrate (mg dl ⁻¹)	2.85 ± 0.35	19.3 ± 2.52*	17.2 ± 1.54*
Plasma insulin (ng ml ⁻¹)	1.23 ± 0.34	0.18 ± 0.03*	0.20 ± 0.05*
Cumulative food intake during refeeding (g chow per mouse)			
2 h	0.39 ± 0.01	1.06 ± 0.05*	1.00 ± 0.05*
4 h	1.20 ± 0.01	1.85 ± 0.05*	1.64 ± 0.03*†
6 h	1.52 ± 0.03	2.32 ± 0.05*	2.43 ± 0.06*
12 h	3.06 ± 0.12	3.83 ± 0.05*	3.82 ± 0.04*
24 h	3.77 ± 0.18	4.50 ± 0.08*	4.43 ± 0.10*
48 h	7.28 ± 0.38	9.26 ± 0.06*	9.20 ± 0.07*

Male C57BL mice (Jackson Laboratory), were housed in plastic cages under constant environmental conditions, a 12-h dark (18:00–06:00h) and 12-h light (06:00–18:00h) cycle, allowed free access to chow and water, and mice were handled twice daily at 09:00 and 18:00 h. At 7 weeks of age, the mice were divided into three groups. After recording body weight and food intake during the dark and light cycles, one group was allowed continued access to chow, the second was deprived of chow for 48 h and received twice daily intraperitoneal (IP) injections of recombinant mouse leptin (Eli Lilly), 1 µg per g body weight, at 09:00 and 18:00 h, and third was fasted for 48 h and treated with saline. Twelve h after the last dose of leptin, the mice were killed by decapitation and trunkal blood obtained for measurement of blood glucose and plasma β-hydroxybutyrate with enzyme assays (Sigma), and insulin by radioimmunoassay (Linco Research). Data are means ± s.e.m., $n = 10$ –13 per group. To determine the effect of leptin treatment on refeeding, the weight of chow consumed by fasted mice was measured 2, 4, 6, 12, 24 and 48 h after the last dose of leptin or saline vehicle at 18:00 h (at the beginning of the dark cycle). Data are means ± s.e.m., $n = 12$ per group. * $P < 0.05$ compared with fed mice; † $P < 0.05$ compared with fasted mice by ANOVA and Fisher PLSD.

decreased the fall in thyroxine to $0.42 \pm 0.05 \mu\text{g dl}^{-1}$ (Table 2). Mean plasma corticosterone and adrenocorticotrophic hormone (ACTH) in fed mice, measured at 11:00 h during the light cycle, were $162 \pm 25 \text{ ng ml}^{-1}$ and $46.7 \pm 4.7 \text{ pg ml}^{-1}$ respectively, and fasting increased corticosterone to $340 \pm 24 \text{ ng ml}^{-1}$ and ACTH to $103 \pm 10.8 \text{ pg ml}^{-1}$ (Table 2). Leptin blunted the rise in corticosterone to $238 \pm 27 \text{ ng ml}^{-1}$ and ACTH to $47.3 \pm 5.1 \text{ pg ml}^{-1}$. Leptin may therefore be involved in mediating the diurnal pattern of the hypothalamic–pituitary–adrenal (HPA) axis.

A diurnal variation of circulating leptin was seen in male mice fed *ad libitum* (Fig. 2b). Leptin concentration was $3.35 \pm 0.62 \text{ ng ml}^{-1}$ at 08:00h, and declined during the light cycle, reaching a nadir early in the dark cycle at 20:00h. Concentration then increased during the dark cycle, coincident with the period of maximum food intake, and reached a peak of $4.2 \pm 0.08 \text{ ng ml}^{-1}$ at 04:00h at the end of the dark cycle. There was a reciprocal temporal relationship between circulating leptin and corticosterone. Serum corticosterone was lowest ($46.3 \pm 9.7 \text{ ng ml}^{-1}$) at the end of the dark cycle, and increased during the light cycle, peaking at $172.3 \pm 31.8 \text{ ng ml}^{-1}$ at 24:00h during the dark cycle. A role of leptin in regulating the diurnal pattern of corticosterone in mice is suggested.

The effects of leptin suggest regulation at the level of the hypothalamus. Leptin receptors have been localized to the choroid plexus and hypothalamus^{17,18} and neuropeptide Y (NPY) is a potential mediator of its central effects⁴. Hypothalamic NPY increases with starvation¹⁹, and intracerebroventricular NPY inhibits gonadotropins and sex steroids²⁰ and activates the HPA axis⁹. The amount of hypothalamic NPY messenger RNA doubled after 48 h of starvation, and was decreased 12 h after leptin treatment (Table 2), consistent with the idea that effects of starvation are mediated by leptin regulation of NPY. A similar role for NPY has been implicated in the

action of leptin in reversing the phenotype of *ob/ob* mice, which also exhibit reproductive failure and increased corticosterone levels responsive to chronic leptin treatment^{4,21}. Although leptin suppressed NPY and exerted neuroendocrine actions during starvation, its effect on recommencement of feeding was minimal, suggesting that factors independent of NPY play a role in compensatory feeding after starvation, as suggested by observations in NPY-knockout mice²².

Several of the responses to starvation were only partially reduced by leptin. Although leptin levels in treated animals

TABLE 2 Effect of fasting and leptin treatment on hormone levels and hypothalamic NPY mRNA

	Fed	Fasted	Fasted + leptin
Testosterone (ng ml ⁻¹)	1.09 ± 0.18	0.12 ± 0.02*	0.45 ± 0.04*†
LH (ng ml ⁻¹)	1.29 ± 0.40	< 0.36	0.52 ± 0.07*†
Thyroxine (µg dl ⁻¹)	0.85 ± 0.05	0.18 ± 0.02*	0.42 ± 0.05*†
Corticosterone (ng ml ⁻¹)	162 ± 24.8	340 ± 24.4*	238 ± 27.1*†
ACTH (pg ml ⁻¹)	46.7 ± 4.7	103 ± 10.8*	47.3 ± 5.1†
NPY mRNA (arbitrary densitometric units)	9.5 ± 0.8	19.5 ± 2.1*	12.9 ± 1.6†

Effect of fasting and recombinant mouse leptin on pituitary–gonadal and pituitary–adrenal hormones and thyroxine. See Table 1 for details of treatment. Testosterone and thyroxine were measured by radioimmunoassay (DPC, Los Angeles), serum luteinizing hormone (LH) by a modification of the rat LH radioimmunoassay using mouse standards²³. Plasma ACTH and corticosterone were measured by radioimmunoassay kits (DPC and ICN Pharmaceuticals, respectively) in mice killed at 11:00h during the light cycle. Data are means ± s.e.m., $n = 10$ –13 per group for testosterone, ACTH, corticosterone and thyroxine. LH was detectable ($>0.36 \text{ ng ml}^{-1}$) by single determination in 8 of 13 fed mice, undetectable in all 13 fasted saline-treated mice, and detectable in 6 of 13 fasted + leptin-treated mice. Total RNA was prepared from individual hypothalami, $n = 3$ per group, and analysed by northern blot hybridization with a complementary RNA probe to mouse NPY as described previously²⁴. * $P < 0.05$ compared with fed controls; † $P < 0.05$ compared with fasted mice by ANOVA and Fisher PLSD for testosterone, corticosterone, ACTH and thyroxine; For LH, * $P < 0.05$ (t-test) compared with fed controls.

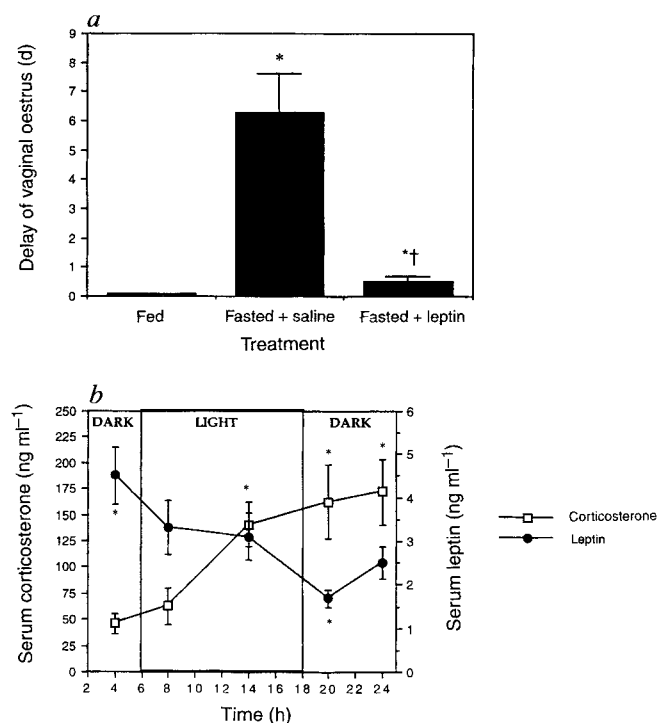


FIG. 2 *a*, Effect of fasting and leptin treatment on the oestrus cycle. Female C57BJ mice were housed singly in plastic cages under ambient conditions, with a 12 h light (06:00–18:00) and 12 h dark (18:00–06:00) cycle, and free access to chow and water. Daily vaginal smears were obtained from age 8 weeks for the duration of 3 oestrus cycles. Thirty females with regular 4–5 day oestrus cycles were assigned to 3 treatment groups (10 per group). One group was fed *ad libitum*, and the others were fasted for 48 h during dioestrus, and treated with twice daily i.p. injections of recombinant mouse leptin, 1 μ g per g body weight or saline. Body weight decreased by 15%, from 23.0 ± 0.9 to 19.6 ± 0.9 gm, and leptin treatment did not alter this. Twelve h after the last injection, all mice were allowed free access to food. Body weight was regained to 98% of control after 24 hours and fully restored after 48 hours. Daily vaginal smears were obtained, each female serving as its own control. The delay of vaginal oestrus was determined as the difference between the length of the cycle (oestrus to oestrus) before and after treatment¹¹. Data are means $\pm$ s.e.m., $n = 10$ per group. * $P < 0.05$ compared with fed controls; † $P < 0.05$ compared with fasted mice by ANOVA and Fisher PSID. *b*, Diurnal variation of serum leptin and corticosterone. Male C57BL mice were used (Table 1), and handling was restricted to cage cleaning. Seventy per cent of food intake (2.78 ± 0.18 g chow per mouse) occurred during the dark cycle, and 30% (1.19 ± 0.10 g chow per mouse) during the light cycle. Groups of mice ($n = 5$) were killed by decapitation at 04:00, 08:00, 14:00, 20:00 and 24:00 h. Serum corticosterone and leptin were measured by radioimmunoassay (ICN and Linco, respectively). Data are means $\pm$ s.e.m., * $P < 0.05$ compared with 08:00 h by ANOVA and Fisher PSID.

were similar to those of fed controls, recombinant leptin may be less potent, or the leptin radioimmunoassay may overestimate bioactive leptin. Alternatively, as falling insulin may mediate adaptation to starvation through regulation of hypothalamic NPY⁹, leptin and insulin may cooperate to regulate aspects of the neuroendocrine response to starvation.

In an environment where periodic limitations of food availability, rather than continuous access, is common, the ability to adapt to starvation is fundamentally important to survival of the species. These studies show that falling leptin concentration is a critical signal that initiates the neuroendocrine response to starvation, including limiting procreation, decreasing thyroid thermogenesis, and increasing secretion of stress steroids, which together are likely to have survival value during prolonged nutritional deprivation. Given the high prevalence of apparent leptin resis-

tance in obese rodents^{5,6} and humans^{6,7}, the physiological response to decreasing leptin concentration with starvation may be the dominant role of this hormone. □

Received 29 April; accepted 30 May 1996.

1. Zhang, Y. et al. *Nature* **372**, 425–432 (1994).
2. Halaas, J. L. et al. *Science* **269**, 543–546 (1995).
3. Campfield, L. A. et al. *Science* **269**, 546–549 (1995).
4. Stephens, T. W. et al. *Nature* **377**, 530–532 (1995).
5. Fredenrich, R. C. et al. *Nature Med.* **1**, 1311–1314 (1995).
6. Maffei, M. et al. *Nature Med.* **1**, 1155–1161 (1995).
7. Considine, R. V. et al. *N. Engl. J. Med.* **334**, 292–295 (1996).
8. Cahill, J. G. et al. *J. clin. Invest.* **45**, 1751–1769 (1966).
9. Schwartz, M. W., Dallman, M. F. & Woods, S. C. *Am. J. Physiol.* **269**, R949–R957 (1995).
10. Dallman, M. F. et al. *Front. Neuroendocr.* **14**, 303–347 (1993).
11. Bronson, F. H. & Marsteller, F. *Biol. Reprod.* **33**, 660–667 (1985).
12. Connors, J. M. et al. *Endocrinology* **117**, 900–906 (1985).
13. Fredenrich, R. C. et al. *J. clin. Invest.* **96**, 1658–1663 (1995).
14. Bray, G. A., Fisler, J. & York, D. A. *Front. Neuroendocr.* **14**, 128–181 (1990).
15. Schreihöfer, D. A. & Cameron, J. L. *Endocrinology* **132**, 1890–1897 (1993).
16. Bates, G. M. *Clin. Obstet. Gynec.* **28**, 632–644 (1985).
17. Tartaglia, L. A. et al. *Cell* **83**, 1263–1271 (1995).
18. Lee, G.-H. et al. *Nature* **379**, 632–635 (1996).
19. White, J. D. & Kershaw, M. *Molec. cell. Neurosci.* **1**, 41–48 (1989).
20. Pierroz, D. Q. et al. *Endocrinology* **137**, 3–12 (1996).
21. Chehab, F. F., Lim, M. E. & Lu, R. *Nature Genet.* **12**, 318–320 (1996).
22. Erickson, J. C., Clegg, K. E. & Palminter, R. D. *Nature* **381**, 415–418 (1996).
23. Cooper, R. L. et al. *Endocrinology* **114**, 391–396 (1984).
24. Qu, D. et al. *Nature* **380**, 243–247 (1996).

ACKNOWLEDGEMENTS. We thank J. Hurwitz for technical assistance and Eli Lilly and Co. for supplying the recombinant leptin. This work was supported by a grant from the NIH to J.S.F. and ADA to E.M.F.

CORRESPONDENCE and requests for materials should be addressed to J.S.F. (e-mail: jfler@bih.harvard.edu).

Consolidation in human motor memory

Thomas Brashers-Krug*[‡], Reza Shadmehr† & Emilio Bizzi*

* Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

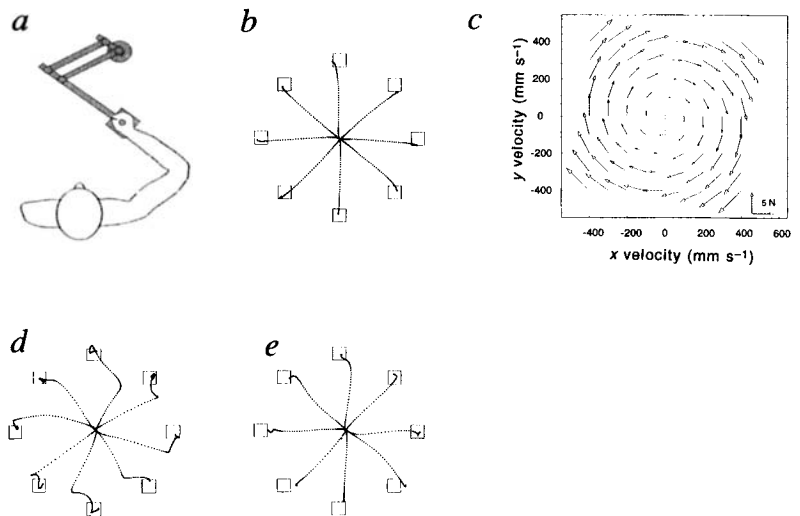
† Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21205, USA

LEARNING a motor skill sets in motion neural processes that continue to evolve after practice has ended, a phenomenon known as consolidation^{1–4}. Here we present psychophysical evidence for this, and show that consolidation of a motor skill was disrupted when a second motor task was learned immediately after the first. There was no disruption if four hours elapsed between learning the two motor skills, with consolidation occurring gradually over this period. Previous studies in humans and other primates have found this time-dependent disruption of consolidation only in explicit memory tasks^{5–12}, which rely on brain structures in the medial temporal lobe^{9,13,14}. Our results indicate that motor memories, which do not depend on the medial temporal lobe^{8,15}, can be transformed by a similar process of consolidation. By extending the phenomenon of consolidation to motor memory, our results indicate that distinct neural systems share similar characteristics when encoding and storing new information.

Subjects moved the handle of a two-link planar manipulator¹⁶ (Fig. 1*a*) to guide a cursor to a series of 192 targets (one target set) that appeared one at a time on a computer monitor mounted above the manipulator (Fig. 1*b*). On the first day of testing (day 1), after baseline trajectories were recorded (Fig. 1*b*),

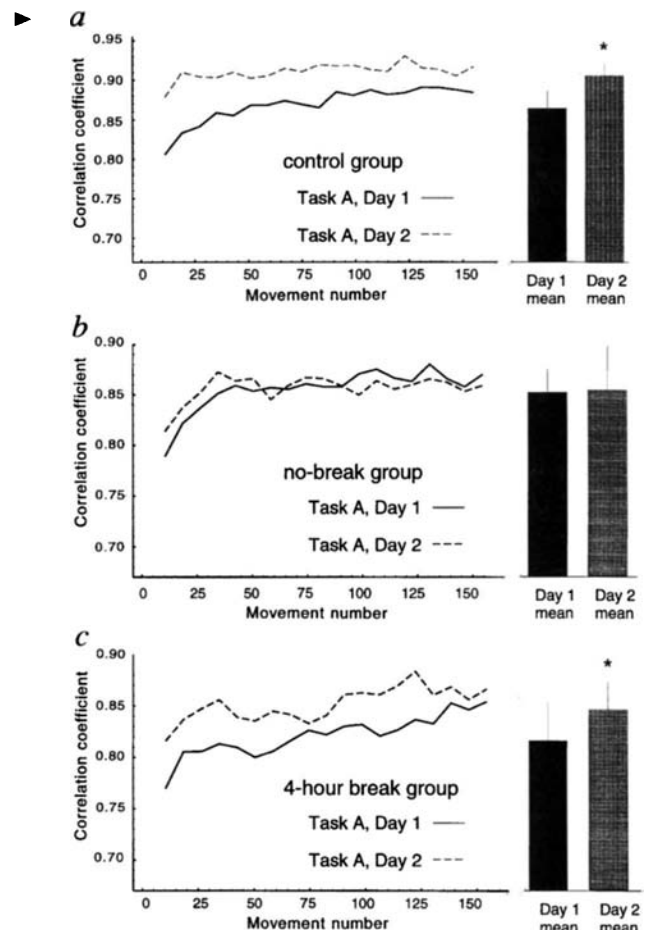
[‡] Present address: Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland 21287, USA.

FIG. 1 We trained 70 naive, right-handed subjects between the ages of 18 and 35 years in a motor learning task¹⁶ (5 subjects were excluded from final analysis because they failed to follow the instructions). Subjects learned to make reaching movements while interacting with a force-producing manipulandum. **a**, Overhead view of the experimental setup. The apparatus is a two-joint planar manipulandum powered by two torque motors mounted at its base³⁰. **b**, Typical hand trajectories for one subject. Subjects moved the handle of the manipulandum 10 cm to targets that appeared in one of eight directions: four directions starting from the centre of the monitor (0°, 45°, 90°, 135°) and the four corresponding directions back to the centre from each of those targets (180°, 225°, 270°, 315°). Subjects were instructed to bring the cursor (representing the position of their hand, displayed on a monitor facing the subjects) in a straight line to each target with a movement time of 500 ± 50 ms. The computer generated a distinctive sound if a subject reached the target within the allotted time. A target turned blue if a subject reached it too slowly, red if too quickly. After 5–10 min practice, when subjects could move at the required pace, we recorded the position and velocity of the manipulandum at 100 Hz for each subject while the subject made 12 movements (baseline trajectories) in each of the 8 target directions. There were no perturbing forces during these movements. **c**, Clockwise velocity-dependent forces imposed by the manipulandum plotted as a function of hand velocity. The direction and length of an arrow indicate the direction and magnitude of the forces, respectively, in each location in velocity space. The movements of half of the subjects were perturbed by clockwise forces, half by anticlockwise forces. Forces were calculated



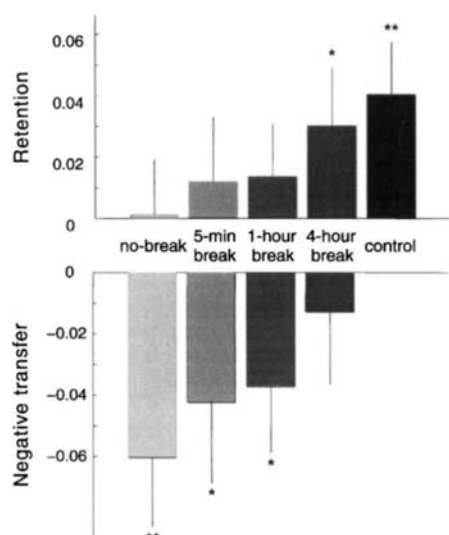
linearly as a function of hand velocity: $f = Bx$. The matrix B was either $\begin{bmatrix} 0 & 13 \\ -13 & 0 \end{bmatrix} \text{ N s m}^{-1}$ or $\begin{bmatrix} 0 & -13 \\ 13 & 0 \end{bmatrix} \text{ N s m}^{-1}$. **d**, Trajectories of one subject when the clockwise perturbing forces were first turned on. **e**, Trajectories of the same subject after 5 min practice in the presence of the clockwise velocity-dependent forces.

FIG. 2 Learning curves for three groups of subjects. The similarity between two trajectories was quantified using a correlation measure that treated each trajectory as a sequence of velocity vectors (x sampled at 10-ms intervals)¹⁶. The value of the correlation coefficient, which depends on both the path and speed of a movement, can vary between -1 and $+1$, with $+1$ indicating that two movements are completely identical. Performance, however, was typically bounded by values of the correlation coefficient between 0.65 and 0.93. Learning curves in task A on days 1 and 2 are shown: **a**, for the control group; **b**, for the no-pause group (task B performed immediately after completion of task A); and **c**, for the 4-hour break group (task B performed 4 hours after completion of task A). The mean performances on days 1 and 2 are shown in bars on the right of each graph (same scale as for learning curves). Mean performance on day 2 was significantly higher than on day 1 for the control and 4-hour break groups, but not for the no-break group. There was no significant difference between the performance of task A on day 1 between these groups. Because subjects were not told in advance of a target set which forces would be applied, the first movement in each target direction was treated as a cue to the pattern of the forces, and was not included in the analysis. (Inclusion of these data, however, did not alter the significance of the comparisons). The asterisk indicates that mean performance on day 2 was significantly higher than on day 1; statistical significance determined with two-tailed t -test, with Bonferroni correction for multiple-planned comparisons.



the manipulandum perturbed subjects' movements with a pattern of velocity-dependent forces during one target set (task A) (Fig. 1c,d). The motor learning task required subjects to compensate for these imposed forces¹⁶. By the end of the target set, subjects were able to guide the cursor accurately to the targets despite the perturbing forces (Fig. 1e). The similarity between a subject's movements in the presence of the forces and his or her baseline movements indicated how well that subject learned to compensate for the forces (Fig. 2).

After the target set, subjects were divided into 6 groups, with 12 subjects in 5 of the groups and 10 in the other. We tested the first



group of subjects (control group) 24 hours later (day 2) with the same forces they had learned in task A. This control group showed both retention of the motor skill and additional learning, performing at a significantly higher level on day 2 than they had on day 1 (Fig. 2).

The second group of subjects was trained on day 1 with a different pattern of forces (task B) immediately after task A (no-break group). In task B, the manipulandum produced forces opposite in direction to those applied during task A. The mean performance of these subjects in task B was significantly less than it had been in task A, an effect known as negative transfer¹⁷. The no-break group was also tested for retention of task A 24 hours later on day 2. In contrast to the control subjects, their mean performance was not significantly better than it had been on the previous day (Fig. 2). This occurred despite the fact that the task provided subjects with continuous visual and kinesthetic feedback on the pattern of forces present. Thus subjects were unable to benefit from their previous training, suggesting that learning task B disrupted the retention of the motor skill that had been learned in task A, a phenomenon known as retrograde interference¹⁷.

To test whether retrograde interference could be caused by any motor task performed immediately after task A, the third group of subjects was tested in a protocol similar to that of the no-break group, the only difference being that, when subjects guided the cursor to the targets in task B, the manipulandum produced no perturbing forces. When tested for retention of task A the following day, these subjects did not demonstrate retrograde interference, performing at a significantly higher level on day 2 than they had on day 1 ($P < 0.05$; data not shown).

We next investigated whether the susceptibility of motor learning to negative transfer and retrograde interference decreased with time. We trained the final three groups of subjects in task B either 5 minutes (5-minute group), one hour (1-hour group), or four hours (4-hour group) after task A on day 1. Both retrograde interference and negative transfer decreased monotonically as the interval between tasks A and B increased (Fig. 3). Although the 5-minute and 1-hour groups performed better in task A on day 2 than they had on day 1, their retention did not approach statistical significance. Both of these groups demonstrated significant negative transfer in task B. When 4 hours passed before task B was learned, however, retention of task A was significant, and there was no significant negative transfer (Figs 2 and 3). Further, there was no difference in the amount of retention between the 4-hour group and the control group ($P > 0.05$); that is, after 4 hours had passed, skill in task A was not disrupted when task B was learned. The initial learning had consolidated.

An alternative explanation for the disappearance of retrograde

FIG. 3 Retention of the task A motor skill and transfer to task B as a function of temporal distance between tasks A and B. Controls did not perform task B. Retention was calculated as the difference between the mean performance in task A on days 2 and 1. Transfer was calculated as the difference between mean performance in task B and task A, both on day 1. An analysis of variance (ANOVA) performed on the retention data indicated a significant difference between the means: $F(4, 48) = 3.05$, $P < 0.05$. An ANOVA performed on the negative transfer data also indicated a significant difference between the means: $F(3, 37) = 2.88$, $P < 0.05$. The asterisk indicates significant difference from 0, $P < 0.05$; two asterisks significant difference from 0, $P < 0.005$; three asterisks indicates significant difference from 0, $P < 0.001$; corrected for multiple comparisons with Bonferroni correction. The concomitant decline of negative transfer and retrograde interference seen in the group data suggests a relation between the two phenomena. This is borne out by the significant correlation ($r = 0.55$, $P < 0.0005$) between negative transfer and retrograde interference in each individual subject.

interference with the passage of time would propose that subjects tended to consider two tasks learned in close succession as a single task: the confusion between the two tasks would diminish as the tasks were separated in time. This explanation places the relevant processes in the cognitive domain, rather than positing a time-dependent change in the way motor memory is stored. Further experiments, however, argued against this interpretation; practising task B 2–3 minutes after task A did not cause retrograde interference (that is, there was still retention of task A) if the subjects had already learned task A on the previous day, and thus had had 24 hours to consolidate their learning (data not shown).

Although several studies have reported the disruption of motor memories^{18,19}, we believe this to be the first evidence that human motor memory, one type of implicit memory, is rapidly transformed with the passage of time, and, in the absence of further practice, from an initial fragile state to a more solid state. Two broad mechanisms may account for this consolidation: either the same synapses that are altered during learning a motor skill are further altered during the changes that lead to the consolidation of that skill; or new synapses are recruited to store the skill in its long-term state^{20–22}. Studies using localized cerebral lesions have suggested that consolidation can occur as a result of a change in the anatomical locus of a memory trace or of changes that rely on the integrity of the medial temporal lobe^{10,23}. These mechanisms, however, seem to take place over a much longer time scale (days to weeks to years) than the 4 hours we observed in our study. Alternatively, the synapses that are altered during learning of a motor skill may be further altered during the changes that lead to the consolidation of that skill^{1,24}. The time course of this latter type of consolidation is consistent with the time course of our results^{1,3,24–29}. □

Received 15 December 1995; accepted 13 May 1996.

1. Bailey, C. H. & Kandel, E.R. In *The Cognitive Neurosciences* (ed. Gazzaniga, M.) 19–36 (MIT Press, Cambridge, MA, 1995).
2. Karni, A. & Sagal, D. *Nature* **365**, 250–252 (1993).
3. Tully, T., Preat, T., Boynton, S. C. & Del Vecchio, M. *Cell* **79**, 35 (1994).
4. Weingartner, H. & Parker, E. (eds) *Memory Consolidation: Psychobiology of Cognition* (L. Erlbaum, Hillsdale, 1984).
5. Muller, G. & Pilzecker, A. Z. *F. Psychol.* **1**, 1–228 (1900).
6. Skaggs, E. *Psychol. Monogr.* **34**, (1925).
7. Sisson, D. E. *J. exp. Psychol.* **25**, 228–233 (1939).
8. Milner, B. *Neuropsychologia* **3**, 317–338 (1965).
9. Squire, L. R. & Alvarez, P. *Curr. Opin. Neurobiol.* **5**, 169–177 (1995).
10. Zola-Morgan, S. & Squire, L. R. *Science* **250**, 288–290 (1990).
11. McGeoch, J. & Irion, A. *Am. J. Psychol.* **43**, 579–592 (1931).
12. Squire, L. R., Slater, P. & Chace, P. *Science* **187**, 77–79 (1975).
13. Malamut, B. L., Saunders, R. C. & Mishkin, M. *Behav. Neurosci.* **98**, 759–769 (1984).
14. Zola-Morgan, S. & Squire, L. R. *Behav. Neurosci.* **99**, 22–34 (1985).
15. Corkin, S. *Neuropsychologia* **6**, 255–265 (1969).
16. Shadmehr, R. & Mussa-Ivaldi, F. A. *J. Neurosci.* **14**, 3208–3224 (1994).

17. McGeoch, J. & Irion, A. *Psychology of Human Learning* 2nd edn, 299–467 (Longmans Green, New York, 1952).
18. Lewis, D., Smith, N. & McAllister, D. J. *exp. Psychol.* **44**, 247–260 (1952).
19. Shea, J. & Titzer, R. J. *Mot. Behav.* **25**, 264–274 (1993).
20. Alvarez, P. & Squire, L. R. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7041–7045 (1994).
21. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. *Psychol. Rev.* **102**, 419–457 (1995).
22. Mishkin, M., Malamut, B. & Bachevalier, J. *Neurobiology of Learning and Memory* (eds Lynch, G., McGaugh, J. L. & Weinberger, N. M.) 65–77 (Guilford, London, 1984).
23. Kim, J. J., Clark, R. E. & Thompson, R. F. *Behav. Neurosci.* **109**, 195–203 (1995).
24. Fujii, S., Saito, K., Miyakawa, H., Ito, K. & Kato, H. *Brain Res.* **555**, 112–122 (1991).
25. Bailey, C. H. & Chen, M. *Science* **220**, 91–93 (1983).
26. Castro-Alamancos, M. A., Donoghue, J. P. & Connors, B. W. J. *Neurosci.* **15**, 5324–5333 (1995).
27. Kimura, A., Caria, M. A., Meis, F. & Asanuma, H. *Neuroreport* **5**, 2372–2376 (1994).
28. Ito, M. *Can. J. Neurol. Sci.* **20**, Suppl. 3, 70–74 (1993).
29. Seeds, N. W., Williams, B. L. & Bickford, P. C. *Science* **270**, 1992–1994 (1995).
30. Faye, I. thesis, MIT (1986).

ACKNOWLEDGEMENTS. We thank E. Fasse, T. Hodgson and N. Hogan for technical support using the MIT manipulandum; M. Smith, K. Thoroughman and K. Akhavan-Toyserkani for technical support using the Johns Hopkins manipulandum; F. Mussa-Ivaldi, S. Corkin and A. Gentile for discussions; and J. Locascio for statistical assistance. This work was supported by grants from the NIH and the ONR to E.B., and from the Whitaker Foundation and the ONR to R.S.

CORRESPONDENCE and requests for materials should be addressed to E.B. (e-mail: emilo@ai.mit.edu).

Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs

Francesco E. Pontieri*, Gianluigi Tanda†, Francesco Orzi*‡ & Gaetano Di Chiara†

* Department of Neuroscience, University 'La Sapienza', Viale dell'Università 30, 00185 Rome, Italy

† Department of Toxicology and CNR Centre for Neuropharmacology, University of Cagliari, Viale A. Diaz 182, 09126 Cagliari, Italy
‡ INM 'Neuromed', Pozzilli (IS), Italy

THE question of whether nicotine, the neuroactive compound of tobacco, is addictive has been open to considerable scientific and public discussion. Although it can serve as a positive reinforcer in several animal species, including man, nicotine is thought to be a weak reinforcer in comparison with addictive drugs such as cocaine and heroin^{1,2}, and has been argued to be habit forming but not addictive^{3,4}. Here we report that intravenous nicotine in the rat, at doses known to maintain self-administration, stimulates local energy metabolism, as measured by 2-deoxyglucose autoradiography, and dopamine transmission, as estimated by brain microdialysis, in the shell of the nucleus accumbens. These neurochemical and metabolic effects are qualitatively similar to

those of other drugs, such as cocaine, amphetamine and morphine, which have strong addictive properties^{5–7}. Our results provide functional and neurochemical evidence that there are specific neurobiological commonalities between nicotine and addictive drugs.

Dopamine neurotransmission in the mesolimbic system, and particularly in the nucleus accumbens, is currently recognized as a critical target of drugs of abuse^{8–10}. Indeed most if not all drugs abused by humans stimulate dopamine transmission in the nucleus accumbens¹¹, a property that has been related to their addictive properties^{8–10}.

The nucleus accumbens is subdivided into a ventromedial 'shell' and a dorsolateral 'core'^{12,13}. The shell is thought to be involved in the integration and expression of emotions, through its projections to the extended amygdala, lateral hypothalamus and central grey matter, and the core is thought to be involved in somato-motor functions^{12,13}.

Several addictive substances such as cocaine, amphetamine and morphine have previously been administered intravenously to freely moving rats at doses that sustain self-administration, and have been shown to increase preferentially or selectively the levels of extracellular dopamine⁷ and energy metabolism^{5,6} in the shell of the nucleus accumbens. Enhanced energy metabolism and increased dopamine transmission in the shell may therefore represent distinctive neurobiological markers of the addictive potential of drugs independently from their specific mechanism of action.

We decided to investigate whether nicotine produces neurochemical effects in the shell that resemble those of typically addictive drugs. Therefore, we studied the effects of nicotine in

TABLE 1 Effects of nicotine on cerebral glucose utilization

	Nicotine dose (mg per kg)		
	0 (saline control) (n = 4)	0.025 (n = 4)	0.050 (n = 4)
Brain area			
Nucleus accumbens shell	80 ± 4	92 ± 5	98 ± 4*
Nucleus accumbens core	85 ± 4	83 ± 3	87 ± 5
Ventral tegmental area	61 ± 2	62 ± 3	65 ± 3
Medial prefrontal cortex	75 ± 7	75 ± 4	71 ± 4
Caudate-putamen (dorsolateral)	109 ± 11	102 ± 6	109 ± 7
Caudate-putamen (dorsomedial)	110 ± 14	102 ± 6	111 ± 7
Caudate-putamen (ventral)	93 ± 8	90 ± 7	89 ± 4
Globus pallidus (dorsal)	60 ± 6	55 ± 7	54 ± 4
Globus pallidus (ventral)	53 ± 5	53 ± 5	52 ± 4
Central amygdala	52 ± 4	46 ± 3	47 ± 3
Basolateral amygdala	88 ± 7	88 ± 6	92 ± 10
Superior colliculus (external)	82 ± 5	94 ± 6	101 ± 10
Lateral geniculate body	89 ± 8	91 ± 5	100 ± 16

Data represent means ± s.e.m. (number of rats in parentheses). Asterisks indicate significant variations with respect to values measured in the control group ($P < 0.05$, one-way analysis of variance, followed by Tukey's *t*-test for multiple comparison). Nicotine, which was administered intravenously to freely moving rats, failed to modify cerebral energy metabolism in the remaining structures: anterior cingulate cortex, sensorimotor cortex, auditory cortex, visual cortex, hippocampus, cerebellar cortex, septal nuclei, lateral hypothalamus, subthalamic nucleus, substantia nigra, thalamic nuclei, habenula, medial geniculate body, inferior colliculus, pontine grey, and corpus callosum.

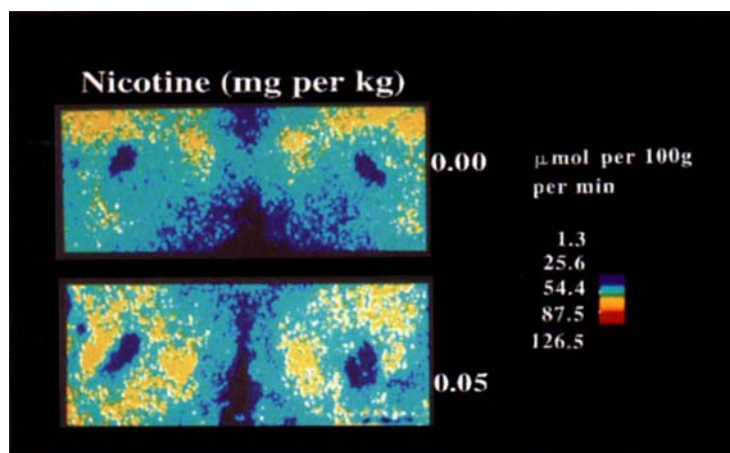
the rat, administered intravenously at unit doses corresponding to those that sustain self-administration^{14,15}, on local cerebral glucose utilization, as measured by 2-[¹⁴C]deoxyglucose autoradiography¹⁶, and extracellular dopamine, as estimated by brain microdialysis^{17,18} using concentric probes aimed at the shell and at the core of the nucleus accumbens⁷.

FIG. 1 Effects of intravenous administration of nicotine on cerebral glucose utilization in the rat nucleus accumbens. Colour-coded transformations of autoradiographs of coronal brain sections at the level of the nucleus accumbens are from a control, saline-treated rat (top) and a rat treated with 0.05 mg per kg nicotine, free base (bottom), in which colours represent rates of glucose utilization. Intravenous administration of nicotine increases energy metabolism in the shell of the nucleus accumbens, without affecting values in the core.

METHODS. Male Sprague-Dawley rats, 280–320 g, were used. Femoral vessels were catheterized under halothane anaesthesia²⁷, after which the rats were allowed at least 3 h to recover before administration of either nicotine or saline ($n = 4$ in each group). The 2-[¹⁴C]deoxyglucose procedure was then begun immediately by injecting an intravenous pulse of 2-[¹⁴C]deoxyglucose (100 μ Ci per kg, specific activity 50–55 mCi mmol⁻¹; Amersham). In other respects the procedure was performed as described¹⁶. Briefly, arterial blood samples were collected over 45 min, centrifuged, and the plasma glucose and 2-[¹⁴C]deoxyglucose concentrations measured⁹. The rats were killed by sodium pentobarbital injection, and brains were removed, frozen and stored at -80°C . Cryostatic coronal sections (20 μ m) were thaw-mounted on glass coverslips, dried at $+60^{\circ}\text{C}$ and autoradiographed on X-ray films (Kodak) with [¹⁴C]methylmethacrylate standards (Amersham). Autoradiographs were analysed by quantitative densitometry using a computerized image-processing system (MCID Imaging Research, Canada). Optical density measurements were

Nicotine, at unit doses of 0.025 and 0.05 mg per kg (free base), induced mild arousal, indicated by sniffing at the air and episodic locomotion. These effects started immediately after administration, and lasted ~ 10 –15 min, depending on the dose.

At a dose of 0.050 mg per kg, nicotine led to increased energy metabolism (as compared to saline controls) only in the shell of



made in at least 5 consecutive sections. Tissue tracer concentrations were determined and a calibration curve obtained from densitometric analysis of autoradiographs of the calibrated standards. Rates of cerebral glucose utilization were calculated, and time courses of arterial plasma glucose and 2-[¹⁴C]deoxyglucose concentrations deduced¹⁶.

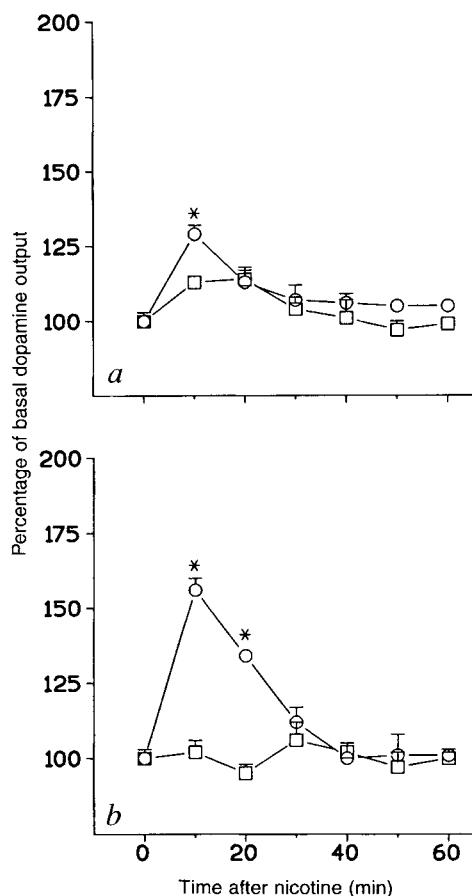


FIG. 2 Effects of intravenous administration of nicotine on dopamine output in dialysates from the shell (circles) and core (squares) subdivisions of the nucleus accumbens in the rat. Percentage variations in dopamine output were produced by nicotine (free base) at 0.25 mg per kg (a) and 0.050 mg per kg (b). Post-hoc analysis indicated a significant increase over basal values in the shell at 10 and 20 min after administration of 0.05 mg per kg nicotine, and at 10 min after administration of 0.025 mg per kg nicotine (asterisk indicates $P < 0.05$ in the shell compared to basal and to the corresponding values in the core).

METHODS. Concentric dialysis probes were prepared with AN 69 fibres (Hospal Dasco, Italy) as described⁷. Rats were anaesthetized with ketamine (100 mg per kg) and placed in a stereotaxic apparatus. The skull was exposed, and a small hole drilled to expose the dura on each side. Each rat was implanted with one dialysis probe on each side (Fig. 3), aimed at the shell on one side and at the core on the other side, according to the rat brain atlas²⁸ (coordinates: shell, $A = +2.0$, $L = 1.2$, $V = 8.0$; core, $A = -1.2$, $L = 2.0$, $V = 7.8$). Anterior, A; lateral, L; ventral, V. A femoral vein was catheterized, and experiments performed 24 h after probes were implanted. Ringer's solution (in mM) (147 NaCl, 2.2 CaCl₂, 4 KCl) was pumped through the dialysis probes at a constant rate of 1 μ l min⁻¹. Dialysate samples (10 μ l) were taken every 10 min and injected without purification into HPLC apparatus equipped with a reverse-phase column (LC-18 DB, 15 cm, particle size 5 μ m; Supelco) and a coulometric detector (ESA, Coulochem II, Bedford, MA) to quantify dopamine. The oxidation electrode of the detector was set at +130 mV, the reduction electrode at -175 mV. The mobile phase (contained) (in mM) 50 NaH₂PO₄, 0.1 Na₂EDTA, 0.5 *n*-octyl sodium sulphate, 18% (v/v) methanol, pH adjusted to 5.5 with Na₂HPO₄. The mobile phase was pumped with a LKB 2150 pump at a flow rate of 1.0 ml min⁻¹. Assay sensitivity for dopamine, 2 fmol per sample.

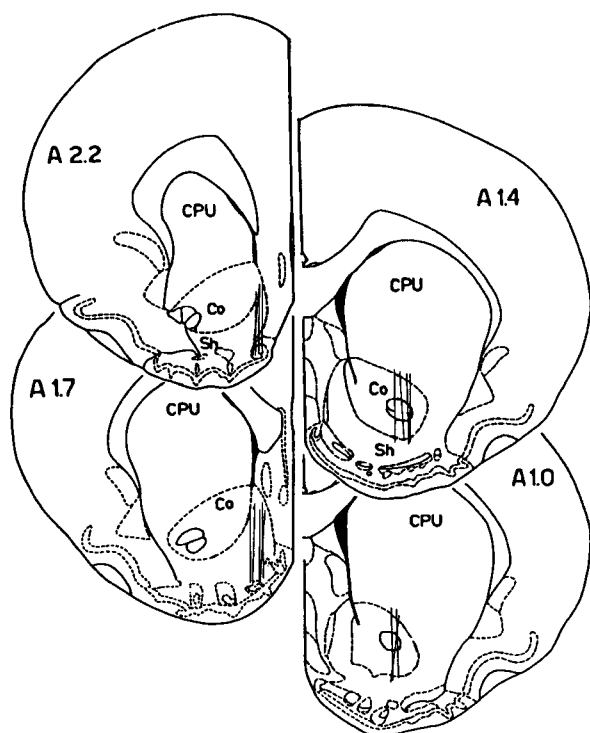


FIG. 3 Localization of the dialysis probes (dialysing portion within the nucleus accumbens and their relationship with shell and core). Abbreviations: A, anterior coordinate²⁸; CPU, candate putamen; Co, Sh, core and shell of the nucleus accumbens.

METHODS. At the end of the experiments, rats were transcardially perfused with 100 ml saline and 500 ml of 4% formaldehyde, 1% calcium acetate, 100 mM NaCl solution. Probes were removed, and brains were cut on a vibratome in serial coronal slices oriented according to the atlas²⁸. Sections were processed for Luxol fast cresyl violet stain and for calbinding immunohistochemistry to identify the location of the probes in relation to the shell and the core of the nucleus.

the nucleus accumbens (Table 1; Fig. 1). Rates of glucose utilization tended to increase in the superficial layer of the superior colliculus and the lateral geniculate body, but the differences with saline controls did not reach statistical significance. In the other 36 brain areas belonging to limbic, motor and sensory systems, nicotine had no effect on energy metabolism. At the lower dose of 0.025 mg per kg, nicotine had no significant effects in any of the brain areas examined.

Administration of nicotine produced dose-dependent increases in dopamine output in dialysates from the shell ($F_{1,20} = 4.268$, $P < 0.05$), with no significant effect in the core ($F_{1,47} = 1.380$, $P = 0.24$) (Figs 2 and 3). *Post-hoc* tests indicated that there was a significant increase in extracellular dopamine in the shell after

10 min for both doses of nicotine (0.025 and 0.05 mg per kg), and after 20 min for 0.050 mg per kg nicotine. There was a significantly larger increase of extracellular dopamine in the shell than in the core after 0.025 mg per kg ($F_{1,54} = 4.233$; *post-hoc*, $P < 0.05$) and 0.05 mg per kg nicotine ($F_{1,47} = 8.260$; *post-hoc*, $P < 0.05$). Because the concentration of dopamine in dialysates reflects extracellular dopamine concentrations, which in turn are an estimate of *in vivo* dopamine transmission^{17,18}, these results indicate that nicotine, at the doses tested, selectively increases dopamine transmission in the shell of the nucleus accumbens.

The selective increase in energy metabolism in the shell is found only after the intravenous administration of a bolus of nicotine, not after intraperitoneal, subcutaneous^{19,20} or continuous intravenous administration²¹. However, the intravenous nicotine pulse used here may more closely approximate the pharmacokinetics of nicotine inhaled in the tobacco smoke. The means of administration has similarly been shown to affect energy metabolism in the highly addictive drug cocaine²².

Our results suggest that the ability of nicotine to increase extracellular dopamine in the nucleus accumbens²³ is the result of a specific activation of mesolimbic dopaminergic neurons projecting to the shell. The nicotine-induced increase in extracellular dopamine in the nucleus accumbens does not undergo tolerance after chronic administration²⁴, suggesting that nicotine-induced activation of dopamine transmission in the shell also occurs after chronic exposure to nicotine.

Stimulation of dopamine transmission and increased glucose metabolism in the shell might be causally related. Indeed, several addictive drugs have been shown to increase both extracellular dopamine and energy metabolism in the shell^{5,6,22}. However, although nicotine activates the firing of dopaminergic neurons in the ventral tegmental area, cocaine and amphetamine actually inhibit this¹⁰. The nicotine-induced increase in energy metabolism in the shell might therefore reflect not only the enhanced energy demand at dopamine terminals associated with increased firing of neurons in the ventral tegmental area, but also the activation of intrinsic non-dopaminergic neurons in the shell, resulting from stimulation of local dopamine receptors by nicotine-induced dopamine release.

Much of the evidence that nicotine is addictive derives from behavioural similarities with drugs that are known to be addictive²⁵. Our results are relevant here because they show that nicotine shares with addictive drugs a discrete neurochemical and functional property, namely, the ability selectively to increase dopamine transmission and energy metabolism in the shell of the nucleus accumbens. This property might itself be critical to the addictive properties of nicotine.

Stimulation of dopamine transmission in the shell of the nucleus accumbens might induce positive motivational states that could lead to the acquisition of incentive properties by stimuli associated with the central effects of nicotine¹⁰. Following this process of motivational learning, stimuli associated with nicotine would become powerful determinants of the compulsive of tobacco²⁶. □

Received 23 February; accepted 17 May 1996.

1. Stolerman, I. P. in *Handbook of Psychopharmacology* (eds Iversen, L. L., Iversen, L. D. & Snyder, S. H.) 421–465 (Plenum, New York, 1987).
2. Swedberg, M. D. B., Henningfield, J. E. & Goldberg, S. R. in *Nicotine Psychopharmacology: Molecular, Cellular and Behavioral Aspects* (eds Wonnacott, S., Russell, M. A. H. & Stolerman, I. P.) 38–76 (Oxford Science, Oxford, 1990).
3. Robinson, J. H. & Pritchards, W. S. *Psychopharmacology* **108**, 397–407 (1992).
4. Warburton, D. M., Revell, A. & Walters, A. C. in *The Pharmacology of Nicotine* (eds Rand, M. & Thureau, K.) 359–373 (IRL, Oxford, 1989).
5. Pontieri, F. E. et al. *NeuroReport* **5**, 2561–2564 (1994).
6. Orzi, F. et al. *Eur. J. Pharmacol.* (in the press).
7. Pontieri, F. E., Tanda, G. & Di Chiara, G. *Proc. natn. Acad. Sci. U.S.A.* **92**, 12304–12308 (1995).
8. Wise, R. A. & Bozarth, M. A. *Psychol. Rev.* **94**, 469–492 (1987).
9. Koob, G. F. *Trends pharmac. Sci.* **13**, 177–184 (1992).
10. Di Chiara, G. *Drug Alcohol Depend.* **38**, 95–121 (1995).
11. Di Chiara, G. & Imperato, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5274–5278 (1988).
12. Alheid, G. F. & Heimer, L. *Neuroscience* **27**, 1–39 (1988).
13. Heimer, L. et al. *Neuroscience* **41**, 89–125 (1991).
14. Corrigan, W. A. & Coen, K. M. *Psychopharmacology* **99**, 473–478 (1989).
15. Corrigan, W. A. in *Effects of Nicotine on Biological Systems* (eds Adklofer, F. & Thureau, K.) 423–

- 432 (Birkhauser, Boston, 1991).
16. Sokoloff, L. et al. *J. Neurochem.* **28**, 897–916 (1977).
17. Ungerstedt, U. in *Measurement of Neurotransmitter Release in Vivo* (ed. Marsden, C. A.) 81–105 (Wiley, Chichester, 1984).
18. Di Chiara, G. *Trends pharmac. Sci.* **11**, 116–121 (1990).
19. London, E. D. et al. *J. Neurosci.* **8**, 3920–3928 (1988).
20. McNamara, D. et al. *J. cerebr. Blood Flow Metab.* **10**, 48–56 (1990).
21. Grunwald, F., Schrock, H. & Kuschinsky, W. *Brain Res.* **400**, 232–238 (1987).
22. Porriño, L. J. *Psychopharmacology* **112**, 343–351 (1993).
23. Imperato, A., Mulas, A. & Di Chiara, G. *Eur. J. Pharmacol.* **132**, 337–338 (1986).
24. Damsma, G., Day, J. & Fibiger, N. C. *Eur. J. Pharmacol.* **168**, 368–371 (1989).
25. Henningfield, J. E. & Heishman, S. J. *Psychopharmacology* **117**, 11–13 (1995).
26. Tiffany, S. T. *Psychol. Rev.* **97**, 147–168 (1990).
27. Crane, A. M. & Porriño, L. J. *Brain Res.* **499**, 87–92 (1989).
28. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates* (Academic, Sydney, 1987).

ACKNOWLEDGEMENTS. These studies were supported by funds from CNR (project FATMA) and from HURST (40% and 60%).

CORRESPONDENCE and requests for materials should be addressed to G.D.C. (e-mail: diptoss@mbx.vol.it).

Excitatory synaptic inputs to spiny stellate cells in cat visual cortex

K. J. Stratford, K. Tarczy-Hornoch, K. A. C. Martin*, N. J. Bannister & J. J. B. Jack

University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

* Institute of Neuroinformatics, ETH and University of Zurich, Gloriastrasse 32, 8006 Zurich, Switzerland

In layer 4 of cat visual cortex, the monocular, concentric receptive fields of thalamic neurons, which relay retinal input to the cortex, are transformed into 'simple' cortical receptive fields that are binocular and selective for the precise orientation, direction of motion, and size of the visual stimulus¹. These properties are thought to arise from the pattern of connections from thalamic neurons¹⁻⁶, although anatomical studies show that most excitatory inputs to layer 4 simple cells are from recurrently connected circuits of cortical neurons⁷⁻⁹. We examined single fibre inputs to spiny stellate neurons in slices of cat visual cortex, and conclude that thalamocortical synapses are powerful and the responses they evoke are unusually invariant for central synapses. However, the responses to intracortical inputs, although less invariant, are strong enough to provide most of the excitation to simple cells *in vivo*. Our results suggest that the recurrent excitatory circuits of cortex may amplify the initial feedforward thalamic signal, subserving dynamic modifications of the functional properties of cortical neurons¹⁰⁻¹².

In cat visual cortex, spiny stellate neurons are the primary target of the lateral geniculate nucleus (LGN) neurons that relay retinal signals to the primary visual cortex^{7,8,13-16}. Their dendrites, which are confined to layer 4, receive synaptic input only from axons forming boutons in layer 4 (refs 9, 14, 17). Excitation comes mainly

from LGN relay neurons, layer 6 pyramidal neurons, and other layer 4 spiny neurons^{9,14-18}. Within layer 4, the axons of layer 6 pyramidal neurons and layer 4 spiny neurons provide about 8 and 5 times as many synaptic contacts, respectively, as LGN neurons⁹. To investigate the properties of these different excitatory synapses, we recorded intracellularly *in vitro* from layer 4a spiny stellate neurons, and evoked fast excitatory postsynaptic potentials (EPSPs) by minimal extracellular stimulation^{19,20} of single afferent axonal fibres. By varying the position of the stimulating electrode we explored the multiple, convergent intracortical and subcortical inputs to a given cell. By using paired recording, we then investigated the excitatory connections between pairs of layer 4 neurons, and between layer 6 and layer 4 neurons.

The minimal-stimulation data are from 21 neurons with spiny stellate morphology (Fig. 1), recorded with γ -aminobutyric acid (GABA_A) receptor-mediated inhibition and NMDA (*N*-methyl-D-aspartate) receptor-mediated excitation blocked, and with 2.5 mM Ca²⁺ and 1 mM Mg²⁺. Multivariate cluster analysis²¹ of 48 inputs showed that they clustered optimally into three classes (Table 1). Class 1 EPSPs (Fig. 1) had large mean amplitudes with low coefficients of variation (CVs), and short response latencies. Stimulation often resulted in failure, alternating on successive trials between the first and second paired pulse; stronger stimulation or longer intervals reduced second-pulse failures, suggesting that these axons are relatively refractory to extracellular stimulation. Paired pulses following successful first pulses were usually depressed in amplitude. In contrast, Class 3 EPSPs (Fig. 2), with high CVs, small mean amplitudes, and longer response latencies, showed paired-pulse facilitation (8 of 12), with a time course consistent with that reported for other cortical neurons²², peaking around 15 ms. Class 2 EPSPs (Fig. 3a-d) showed intermediate CVs and amplitudes, latencies similar to those of class 3 EPSPs, and slight depression or no change for paired pulses.

There were several indications that the three classes of synaptic response might be correlated with the three major anatomical sources of excitation to spiny stellate neurons⁹. Class 1 inputs had the shortest response latencies, despite distant sites of stimulation,

TABLE 1 Summary of statistics for EPSPs of different synaptic classes

	Class 1		Class 2		Class 3	
	Minimal stimulation		Dual recording		Minimal stimulation	
Number	20		14		7	
Amplitude (μ V)	1,970 $\pm$ 1,447 (387-5,887)		1,072 $\pm$ 689 (178-2,205)		334 $\pm$ 217 (123-788)	
CV (%)	8.3 $\pm$ 3.9 (-2.5-16.5)		21.6 $\pm$ 10.1 (9.4-39.9)		47.0 $\pm$ 13.9 (5.5-72.0)	
Rise time (ms)	1.1 $\pm$ 0.5 (0.4-1.8)		1.1 $\pm$ 0.4 (0.8-1.8)		1.4 $\pm$ 0.5 (0.8-2.0)	
Half width (ms)	8.8 $\pm$ 4.7 (3.2-13.2)		10.7 $\pm$ 4.2 (5.4-17.8)		7.8 $\pm$ 2.0 (4.8-10.2)	
Latency (ms)	1.1 $\pm$ 0.3 (0.6-1.8)		1.7 $\pm$ 0.7 (0.8-3.0)		1.7 $\pm$ 0.7 (1.0-3.4)	
AQD	0.17 $\pm$ 0.08 (0.06-0.30)		0.36 $\pm$ 0.07 (0.23-0.45)		0.41 $\pm$ 0.04 (0.32-0.47)	

For each synaptic input the following variables were recorded: amplitude (mean amplitude of successful responses to first stimulus at 1 Hz averaged for all available stimulus trials); CV (defined as $(\sigma_s^2 - \sigma_n^2)^{1/2} / \mu_s$, excluding stimulation failures, and expressed as a percentage (in two cases the EPSP variance was less than that of the noise, and CVs were defined as $-(\sigma_s^2 - \sigma_n^2)^{1/2} / \mu_s$); rise time (time for rising phase of EPSP from 10% to 90% of peak amplitude); half width (time between 50% peak amplitude on rising phase and 50% on falling phase of EPSP); latency (time from onset of stimulus activation to onset of EPSP); and average quadrant deviation (AQD) (see Fig. 1 legend). All figures are shown as mean $\pm$ standard deviation (minimum - maximum). The 48 synaptic inputs from minimal-stimulation experiments were analysed with a program that performed cluster analysis²¹. Using the parameters of coefficient of variation, latency, direction of paired-pulse change (depression or facilitation defined as greater than 10% change), and AQD in assessing the clustering criterion, it was found that allowing three classes gave the optimal categorization of the inputs, as apparent from a marked inflection in the plot of the clustering criterion as a function of the number of groups²¹. The sample numbers do not reflect the true prevalence of each class of input, because we selected positively for the large-amplitude class 1 and class 2 inputs, although class 3 inputs were actually most common. Class 2 and class 3 extracellularly evoked EPSPs are very similar to EPSPs recorded in dual impalements from layer 4 and layer 6 presynaptic neurons, respectively. The tendency of extracellularly evoked inputs in each group to have lower CVs than the corresponding dual-recording group is caused by the requirements of a rigorous minimal-stimulation protocol, which generally biases acceptance in favour of less variable inputs. In dual recordings it was not possible to measure latency, as the EPSP onset was obscured by the presynaptic spike artefact. This may also have affected the rise-time measurement.

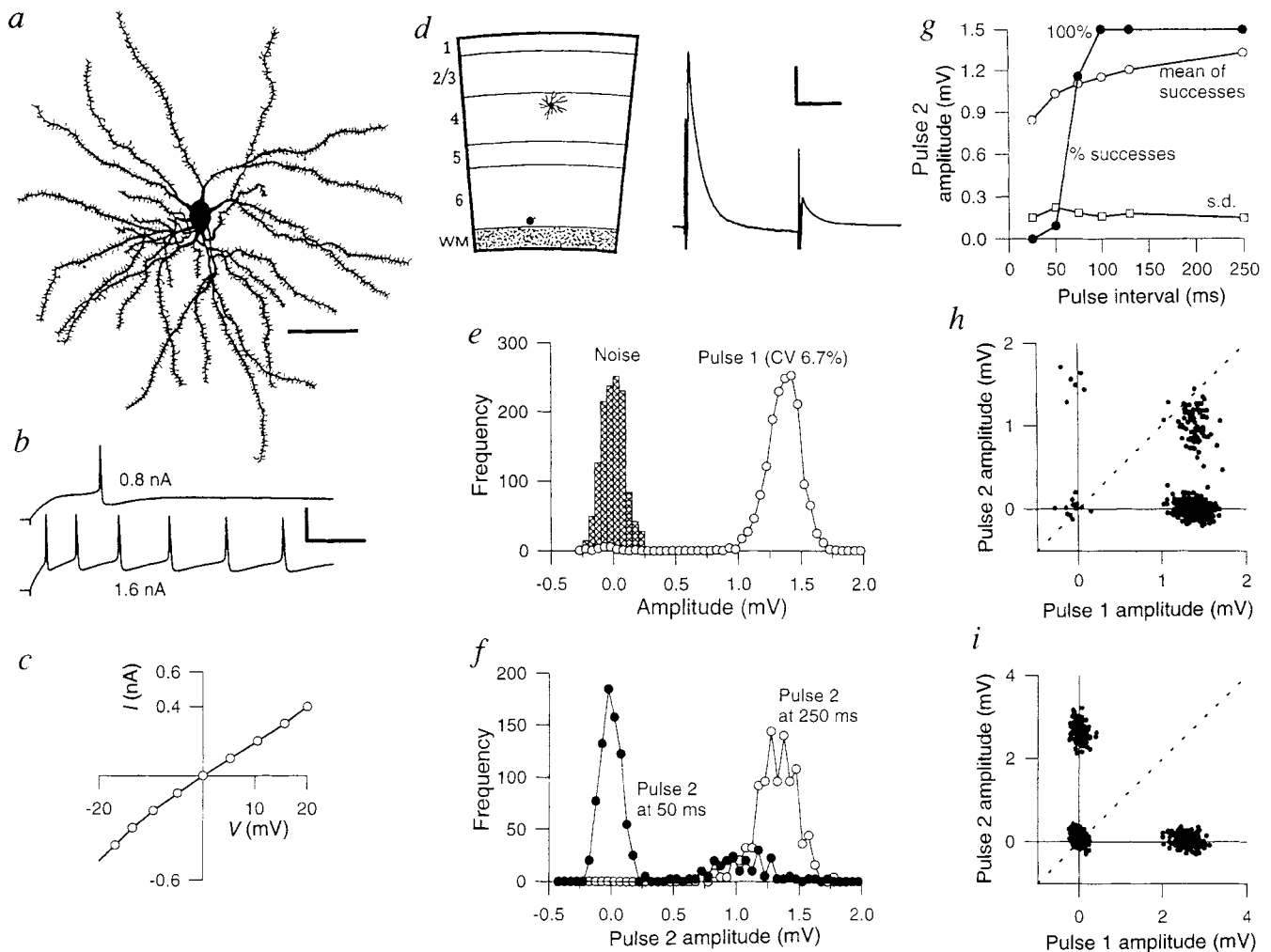


FIG. 1 Class 1 inputs. *a*, Camera lucida reconstruction of a spiny stellate neuron from which the data in *b–h* were recorded. Scale bar, 50 μm . *b*, Firing responses to 0.8 nA and 1.6 nA current pulses. Scale bar, 30 ms, 50 mV. *c*, Current–voltage relation (100-ms current steps; origin –76 mV). Average resting potential, input resistance and membrane time constant of cells (2 M electrodes) were –56 mV, 63 M Ω , and 16 ms (holding potential –80 mV). *d*, Left, positions of layer 4 neuron, and stimulating electrode at border of layer 6 and white matter (WM) (50 μm off-radial towards the LGN) for the input described in *d–h* (1,650 trials recorded). Right, averaged EPSP is of large amplitude; averaged response to second pulse is small owing to stimulation failures. With membrane depolarization (data not shown), amplitudes became smaller (CVs invariant), consistent with passive membrane properties, suggesting that large amplitudes were not from activation of voltage-dependent dendritic Na⁺ or Ca²⁺ channels. Scale bar, 20 ms, 300 μV . *e*, Amplitude histograms of noise and pulse-1 response: $\sigma_n = 82 \mu\text{V}$, $\mu_{s1} = 1.373 \mu\text{V}$, and $\sigma_{s1} = 123 \mu\text{V}$; noise-corrected CV 6.7%. *f*, Pulse-2 response at 50 ms (filled circles) and 250 ms (open circles) paired-pulse intervals; at 50 ms, pulse 2 fails more, and responses to successful stimulation are depressed. *g*, As interpulse interval is shortened, mean amplitude of successful pulse 2 responses decreases gradually and the proportion of failures increases sharply. *h*, Scattergram of pulse-2 response (50 ms delay) against pulse-1 response; depression illustrated by points just below diagonal. *i*, Similar plot for a different class-1 input; all trials show failure on pulse 1 and/or pulse 2. To quantify these scattergram patterns, axes were divided at the mid-range of pulse-1 responses, and the mean distance of points from the centroid of the points

in each quadrant, averaged over four quadrants, was normalized by quadrant size to give an average quadrant deviation score for use in cluster analysis (Table 1).

METHODS. Blocks of visual cortex (area 17) from cats 12–14 weeks old (1.0–1.3 kg) were excised by craniotomy after pentobarbitone-induced anaesthesia (Sagatal, Sigma, 60 mg per kg, intraperitoneal) maintained with intravenous Saffan. Coronal slices (400 μm thick) were maintained in an interface chamber²⁵, with 50 μM D,L-2-amino-5-phosphopentanoic acid (AP5, Sigma) and, for minimal stimulation, 100 μM picrotoxin (Sigma). Voltage recordings of synaptic potentials and responses to current pulses in layer 4a neurons used biocytin-filled sharp intracellular microelectrodes²⁵ (0.5 M CH₃SO₃K filling solution for dual impalements, 2 M for minimal-stimulation experiments). A minimal-stimulation protocol^{19,20} activated single afferent axons with a bipolar wire electrode (20–200 μs voltage pulses): the average evoked response as a function of stimulus intensity shows plateaux separated by step increases as individual fibres are recruited²⁰; stimulus intensity was adjusted to the first plateau, provided this response showed no ‘contamination’ by more than one fibre (very variable latency or rise time). Responses were evoked at 1 Hz, using paired pulses (interval 50 ms or varied systematically where appropriate). In dual recording the presynaptic cell was driven to fire multiple spikes to explore paired-pulse behaviour. Synaptic responses were digitized at 5 kHz, and amplitudes of individual events were measured as differences between averages over short windows at baseline and signal peak²⁵. Fixed slices were resectioned and biocytin visualized for morphological identification and reconstruction.

usually at the border between layer 6 and white matter (16 of 20). They must have arisen from rapidly conducting axons, such as the myelinated thalamic axons. These fibres are thickest, and thus easiest to activate where they enter the grey matter¹⁸, which would explain why class 1 inputs were preferentially evoked there. Severed axons might be prone to prolonged refractoriness and hence to stimulation failures, which were observed only in class 1 inputs. Class 2 inputs could be evoked throughout layers 2–6, but were the only ones activated from layers 2 and 3. Thalamic and layer 6 pyramidal axons rarely project to superficial layers^{7,13}, whereas the axons of layer 4 spiny neurons project heavily into layers 2 and 3 (refs 14–16), and would be easily activated there. Layer 6 pyramidal neurons, unlike thalamic afferents, produce an augmenting response in layer 4 *in vivo*²³, possibly by means of EPSPs showing paired-pulse facilitation, like the class 3 inputs. Indeed, most (9 of 10) class 3 EPSPs were evoked radially in layers 5 and 6, where layer 6 pyramidal axons ascend. All three input classes converged on individual spiny stellate neurons, as predicted by anatomical observations⁹.

We investigated this suggested structure–function correlation directly by examining the intracortical subset of inputs to layer 4 using dual intracellular recording²⁵. We recorded 7 connections from layer 6 pyramids to layer 4 excitatory neurons. These EPSPs

had small amplitudes and large CVs, and showed paired-pulse facilitation (7 of 7), thus closely resembling class 3 minimal stimulation inputs (Table 1 and Fig. 2e–i). Paired recording between excitatory cells in layer 4 yielded 14 connections onto confirmed spiny stellate neurons. Like class 2 EPSPs, these had larger amplitudes and lower CVs than the layer 6 to 4 connections, and showed depression or no change on paired pulses (Table 1 and Fig. 3e–h). The match between minimal stimulation and dual recording data validates the minimal-stimulation technique, and importantly allows us to attribute the physiological characteristics of class 2 and class 3 EPSPs directly to defined intracortical excitatory pathways. Class 1 EPSPs were not seen in any dual-impalement connections ($n = 51$), thus further supporting their suggested extracortical, thalamic origin: *in vitro* axons can be activated by extracellular stimulation, but are unlikely to be impaled.

The putative thalamic EPSPs show a striking combination of large amplitude and low CV, and, along with class 2 EPSPs, contradict the suggestion that cortical connections always evoke small, highly variable responses²⁶. Class 1 CVs imply minimal trial-to-trial variation in the number of neurotransmitter quanta released, as if there were only one release site. However, the EPSPs are large for monoquantal responses. From EPSP time

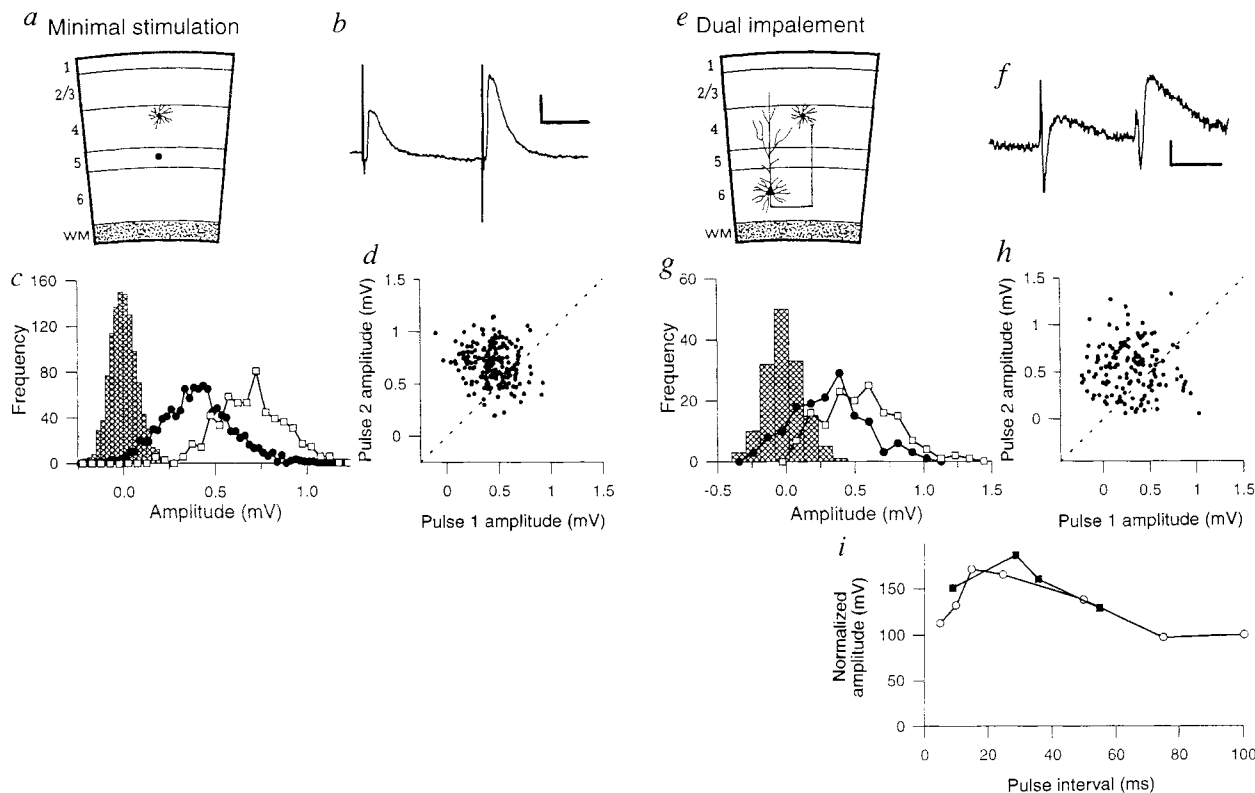


FIG. 2 Class 3 inputs. Characteristics of class 3 synaptic inputs seen with extracellular stimulation (a–d) and dual impalements (e–h). *a*, Schematic positions of neuron (same neuron as in Fig. 1) and stimulating electrode (radial in layer 5) for the input described in b–d and i (1,200 trials recorded). *b*, Averaged EPSP shows small amplitude with paired-pulse facilitation. Scale bar, 20 ms, 200 μ V. *c*, Amplitude histograms of baseline noise (hatched bars) and responses to first (filled circles) and second pulses (open squares). The response to the second stimulus (at 15 ms delay) is facilitated relative to the first ($\sigma_n = 75 \mu$ V, $\mu_{s1} = 402 \mu$ V, $\sigma_{s1} = 187 \mu$ V, $\mu_{s2} = 687 \mu$ V, $\sigma_{s2} = 181 \mu$ V; CV of pulse 1 is 42.6%). *d*, Paired-pulse scattergram showing facilitation (more points above the diagonal than below); the response to each stimulus is highly variable. *e*, Schematic positions of presynaptic layer 6 neuron and postsynaptic spiny stellate neuron for the dual-recording synaptic connection described in f–h (1,300

trials recorded). Presynaptic neurons when recovered were all layer 6 pyramidal neurons; postsynaptic neurons were either spiny stellate neurons or layer 4 pyramidal neurons. *f*, Averaged EPSP shows small amplitude with paired-pulse facilitation. Scale bar, 20 ms, 200 μ V. *g*, Amplitude histograms of baseline noise (hatched bars) and responses to first (filled circles) and second pulses (open squares). The second response (31 ms mean delay) is facilitated relative to the first ($\sigma_n = 133 \mu$ V, $\mu_{s1} = 308 \mu$ V, $\sigma_{s1} = 267 \mu$ V, $\mu_{s2} = 533 \mu$ V, $\sigma_{s2} = 262 \mu$ V; CV of pulse 1 is 75.2%). *h*, Paired-pulse scattergram showing facilitation and variability of response to each stimulus. *i*, Facilitation (100 $\times$ pulse 2/pulse 1) of the minimal stimulation input (open circles) described in a–d as a function of interpulse interval (200 trials per interval); facilitation is maximal (71%) at 15 ms delay. Dual-impalement data (filled squares) from the connection in e–h shows a similar dependence of facilitation on interpulse interval.

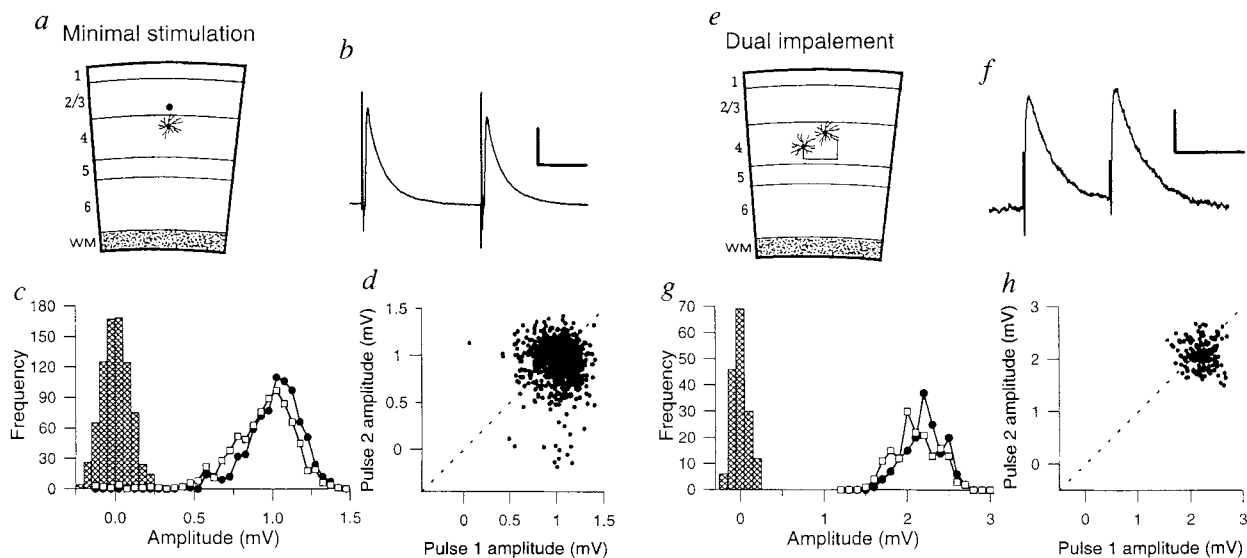


FIG. 3 Class 2 inputs. Characteristics of class 2 synaptic inputs seen with extracellular stimulation (a–d) and dual impalements (e–h). **a**, Schematic positions of neuron (same neuron as in Fig. 1) and stimulating electrode (radial in layer 3) for the input described in b–d (800 trials recorded). **b**, Averaged EPSP shows an intermediate amplitude response to both pulses. Scale bar, 20 ms, 400 μ V. **c**, Amplitude histograms of baseline noise (hatched bars) and responses to first (filled circles) and second pulses (open squares). The response to the second stimulus (at 50 ms delay) is slightly decreased in amplitude and similar in variation to the first ($\sigma_n = 92 \mu$ V, $\mu_{s1} = 1.015 \mu$ V, $\sigma_{s1} = 169 \mu$ V, $\mu_{s2} = 932 \mu$ V, $\sigma_{s2} = 218 \mu$ V; CV of pulse 1 is 13.9%). **d**, Paired-pulse scattergram shows that stimulation rarely fails to evoke a response to both stimuli, and the second response is similar to the first; both responses show variation in amplitude.

e, Schematic positions of presynaptic and postsynaptic neurons in layer 4 for the dual-recording synaptic connection described in f–h (600 trials recorded). Both neurons were morphologically confirmed spiny stellate neurons (somata 55 μ m apart laterally). **f**, Averaged EPSP again shows similar amplitude responses to both pulses. Scale bar, 40 ms, 800 μ V. **g**, Amplitude histograms of baseline noise (hatched bars) and responses to first (filled circles) and second pulses (open squares). The response to the second stimulus (at 54 ms mean delay) is slightly decreased in amplitude and similar in variation to the first ($\sigma_n = 88 \mu$ V, $\mu_{s1} = 2.205 \mu$ V, $\sigma_{s1} = 226 \mu$ V, $\mu_{s2} = 2.097 \mu$ V, $\sigma_{s2} = 259 \mu$ V; CV of pulse 1 is 9.4%). **h**, Paired-pulse scattergram shows the second response is essentially similar to the first.

integrals and cell input resistances²⁷, we calculated the number of channels opening during the response to produce each EPSP, assuming that the postsynaptic receptors of cat and rat visual cortical spiny neurons have similar channel conductances and mean open times, about 9 pS and 1.9 ms, respectively²⁸. Class 1 EPSPs require 428 ± 302 channels ($n = 14$); class 2, 280 ± 171 ($n = 14$); and class 3, 79 ± 39 ($n = 8$). Thus class 1 responses probably involve multiple synaptic contacts, all with extremely high release probability, whereas class 2 and class 3 EPSPs are probably multiquantal with moderately high and low release probabilities, respectively. Inputs with some class 1 characteristics have recently been reported in rat visual cortex²⁰.

There is considerable support for the view that the basic structure of the simple receptive field is established by the pattern of thalamic innervation^{1–6}. However, our data suggest that thalamic excitation by powerful class 1 EPSPs will be amplified by the more numerous synaptic connections from adjacent layer 4 neurons. Additionally, temporally sensitive, potentiating synapses from layer 6 pyramidal neurons, which provide the major excitatory pathway to layer 4, can augment and modulate thalamic signal transmission^{23,24}. These recurrent cortical circuits may have several functions, including contrast gain control¹⁰, extraction of signals from incomplete or noisy data²⁹, and context-dependent receptive-field modifications^{11,12}. □

Received 1 April; accepted 21 May 1996.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962).
- Chapman, B., Zahs, K. R. & Stryker, M. P. *J. Neurosci.* **11**, 1347–1358 (1991).
- Reid, R. C. & Alonso, J.-M. *Nature* **378**, 281–284 (1995).
- Ferster, D., Chung, S. & Wheat, H. *Nature* **380**, 249–252 (1996).
- Hubel, D. H. *Nature* **380**, 197–198 (1996).
- Das, A. *Neuron* **16**, 477–480 (1996).
- Garey, L. J. & Powell, T. P. *Proc. R. Soc. Lond. B* **179**, 41–63 (1971).
- Peters, A. & Payne, B. R. *Cereb. Cortex* **3**, 69–78 (1993).
- Ahmed, B., Anderson, J. A., Douglas, R. J., Martin, K. A. C. & Nelson, J. C. *J. comp. Neurol.* **341**, 39–49 (1994).
- Solar, G. & Freeman, R. D. *Expl. Brain Res.* **46**, 457–461 (1982).
- Gilbert, C. D. & Wiesel, T. N. *Vision Res.* **30**, 1689–1701 (1990).
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. & Davis, J. *Nature* **378**, 492–496 (1995).
- LeVay, S. & Gilbert, C. D. *Brain Res.* **113**, 1–19 (1976).
- Lund, J. S., Henry, G. H., MacQueen, C. L. & Harvey, A. R. *J. comp. Neurol.* **184**, 599–618 (1979).
- Gilbert, C. D. & Wiesel, T. N. *Nature* **280**, 120–125 (1979).
- Martin, K. A. C. & Whitteridge, D. J. *Physiol., Lond.* **353**, 463–504 (1984).
- Anderson, J. A., Douglas, R. J., Martin, K. A. C. & Nelson, J. C. *J. comp. Neurol.* **341**, 25–38 (1994).
- Hirsch, J. J. *Physiol., Lond.* **483**, 183–199 (1995).

- Allen, C. & Stevens, C. F. *Proc. natn. Acad. Sci. U.S.A.* **91**, 10380–10383 (1994).
- Volgushev, M., Voronin, L. L., Chistiakova, M., Artola, A. & Singer, W. *Eur. J. Neurosci.* **7**, 1751–1760 (1995).
- Everitt, B. S. *Cluster Analysis* 3rd edn (Edward Arnold, London, 1993).
- Andreassen, M. & Häblitz, J. J. *J. Neurophysiol.* **72**, 326–336 (1994).
- Ferster, D. & Lindström, S. J. *Physiol., Lond.* **367**, 233–252 (1985).
- Grieve, K. L. & Sillito, A. M. *Expl. Brain Res.* **87**, 521–529 (1991).
- Mason, A., Nicol, A. & Stratford, K. J. *Neurosci.* **11**, 72–84 (1991).
- Stevens, C. R. in *Large-scale Neuronal Theories of the Brain* (eds Koch, C. & Davis, J. L.) 239–255 (MIT Press, Cambridge, MA, 1994).
- Jack, J. J. B., Noble, D. & Tsien, R. W. *Electrical Current Flow in Excitable Cells* (Clarendon, Oxford, 1975).
- Hestrin, H. *Neuron* **11**, 1083–1091 (1993).
- Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C. & Suarez, H. H. *Science* **269**, 981–984 (1995).

ACKNOWLEDGEMENTS. We thank A. Larkman for valuable advice and for reading the manuscript; D. Fleming for animals; and B. Ahmed, J. Anderson, V. Meskenaitė and C. Nelson for morphological advice and help. K.T.-H. is a Wellcome Trust prize student. This work was supported by The Wellcome Trust, The Royal Society and The Medical Research Council.

CORRESPONDENCE and requests for materials should be addressed to K.J.S. (e-mail: kenneth.stratford@physiol.ox.ac.uk).

Identification of the homologous beige and Chediak–Higashi syndrome genes

Maria D. F. S. Barbosa*, Quan A. Nguyen*, Velizar T. Tchernev*, Jennifer A. Ashley†, John C. Detter*, Susan M. Blaydes*, Stephen J. Brandt‡, Dipti Chotai§, Charles Hodgman§, Roberto C. E. Solari§, Michael Lovett† & Stephen F. Kingsmore*

* Departments of Medicine and Pathology, and Center for Mammalian Genetics, University of Florida, Gainesville, Florida 32610, USA

† Department of Biochemistry and the McDermott Center, The University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA

‡ Departments of Medicine and Cell Biology, Vanderbilt University, Nashville, Tennessee 37232, USA

§ Cell Biology, Advanced Technologies and Informatics Units, Glaxo Wellcome Medicines Research Center, Stevenage, Hertfordshire SG1 2NY, UK

VESICULAR transport to and from the lysosome and late endosome is defective in patients with Chediak–Higashi syndrome (CHS) and in mutant beige (*bg*) mice^{1–4}. CHS and *bg* cells have giant, perinuclear vesicles with characteristics of late endosomes and lysosomes that arise from dysregulated homotypic fusion^{3–5}. CHS and *bg* lysosomes also exhibit compartmental missorting of proteins, such as elastase, glucuronidase and cathepsin G^{2,3,6,7}. *Lyst*, a candidate gene for *bg*, was identified by direct complementary DNA selection from a yeast artificial chromosome (YAC) clone containing a 650-kilobase segment of the *bg*-critical region on mouse chromosome 13. *Lyst* is disrupted by a 5-kilobase deletion in *bg*¹¹ mice, and *Lyst* messenger RNA is markedly reduced in *bg*²¹ homozygotes. The homologous human gene, *LYST*, is highly conserved with mouse *Lyst*, and contains a frame-shift mutation at nucleotides 117–118 of the coding domain in a CHS patient. Thus *bg* mice and human CHS patients have homologous disorders associated with *Lyst* mutations. *Lyst* encodes a protein with a carboxy-terminal prenylation motif and multiple potential phosphorylation sites. *Lyst* protein is predicted to form extended helical domains, and has a region of sequence similar to stathmin, a coiled-coil phosphoprotein thought to act as a relay integrating cellular signal response coupling^{8–10}.

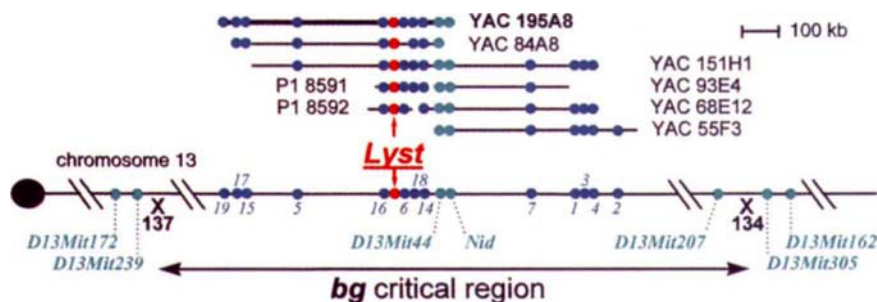
FIG. 1 Genetic and physical map of the *bg* non-recombinant interval on mouse chromosome 13 showing the location of *Lyst*. Mouse chromosome 13 is shown by a horizontal line with the centromere on the left. The *bg* critical region is delineated by chromosome crossovers (denoted with an X) in animals 134 and 137 of an intersub-specific mouse backcross¹² [C57BL/6J – *bg*¹ × (C57BL/6J – *bg*¹ × CAST/EiJ)F₁]. Microsatellite markers D13Mit172 and D13Mit239 flank *bg* proximally; D13Mit162 and D13Mit305 lie distal to *bg* (indicated by turquoise circles). YAC and P1 clones identified by PCR screening^{25,26} with oligonucleotides corresponding to *Nid* or D13Sf13 are shown above the chromosome. Novel sequence-tagged sites (STSs, indicated by dark blue circles), generated by inverse repetitive element PCR or direct cDNA selection¹³, were used to order clones within the contiguous array. Novel mouse chromosome 13 STSs are numbered 1–18, corre-

We have previously localized the *bg* locus within a 0.24-centimorgan interval on mouse chromosome 13, and isolated contiguous arrays of YACs that cover 2,400 kb of this interval^{11–13}. Candidate cDNAs for *bg* were isolated from YAC 195A8, which contains 650 kb of the *bg* non-recombinant interval using direct cDNA selection with mouse spleen cDNA¹³ (Fig. 1). Of 56 candidate cDNA clones analysed from a direct-selection experiment, evidence for causality in *bg* was found in one (see below), and this gene was designated *Lyst* (lysosomal trafficking regulator). As this clone was 132 nucleotides long, additional *Lyst* sequences were sought by screening three mouse cDNA libraries and performing polymerase chain reaction (PCR) amplification of cDNA ends¹⁴. Ten overlapping *Lyst* clones were identified, representing ~7 kb (Genbank accession number, L77889). These were physically assigned to mouse chromosome 13 with pulsed field gel electrophoresis (PFGE) Southern blots, confirming that they were all derived from a single gene (mouse genome database accession number, MGD-PMEX-14). The *Lyst* probes identified the same polymorphic PFGE restriction fragments as *nidogen* (*Nid*)¹², indicating that *Lyst* and *Nid* are clustered within 650 kb. *Lyst* was also mapped genetically in 504 [C57BL/6J-*bg*¹ × (C57BL/6J-*bg*¹ × CAST/EiJ)F₁] backcross mice¹² by means of three *TaqI* restriction fragment length polymorphisms (RFLPs). The *Lyst* RFLPs cosegregated with *bg* (and *Nid*), confirming their colocalization on proximal mouse chromosome 13 (MGD accession number, MGD-CREX-615).

Evidence for *Lyst* mutations was found in two *bg* alleles. A 5-kb genomic deletion that contained the 3' end of *Lyst* exon β, and exons γ and δ, was identified in *bg*¹¹ DNA (Fig. 2). The *bg*¹¹ deletion corresponds to loss of ~400 internal amino acids of the predicted *Lyst* peptide. Furthermore, whereas the 5' end of the *bg*¹¹ deletion occurs within *Lyst* exon β, the 3' end is intronic. Therefore the truncated *Lyst* mRNA in *bg*¹¹ mice is also anticipated to splice incorrectly, terminate prematurely, and lack polyadenylation.

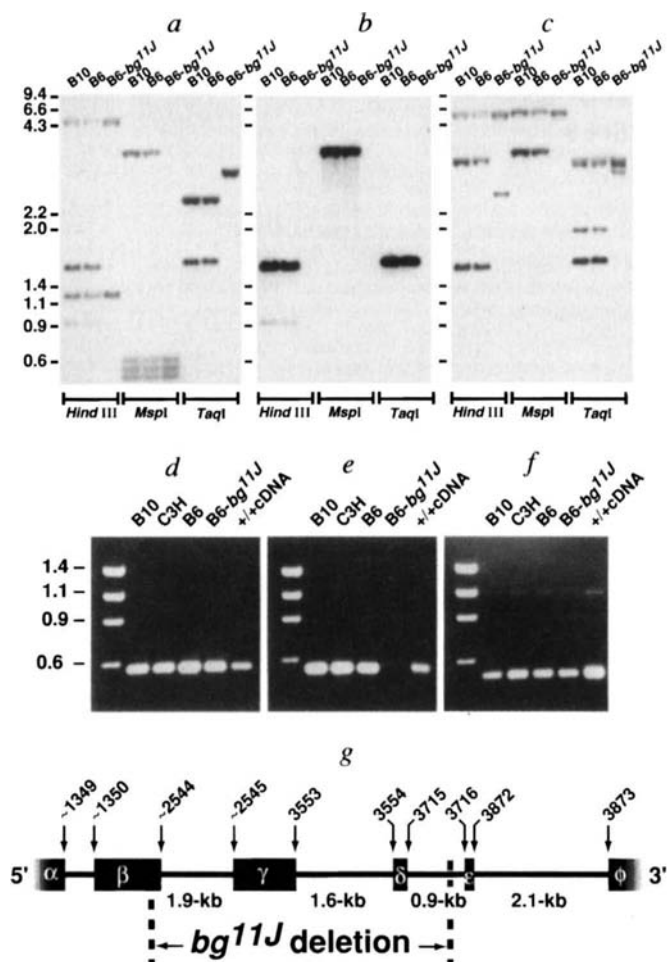
Quantitative reverse transcription (RT)–PCR demonstrated a moderate decrease in *Lyst* mRNA in *bg* and *bg*¹ liver, and a gross reduction in *bg*²¹ (*Lyst* ΔOD after normalization for β-actin mRNA: +/+ , 1.00; *bg*²¹/*bg*²¹, 0.19; *bg*/*bg*, 0.28; *bg*¹/*bg*¹, 0.40). A commensurate reduction in *bg*²¹ transcript abundance was noted by using several primer pairs derived from different regions of the *Lyst* cDNA. Aberrant *Lyst* RT–PCR products were not observed. The particularly striking (more than fivefold) reduction in *Lyst* expression evident in *bg*²¹ homozygotes suggested the existence of a mutation in *bg*²¹ *Lyst* that results in decreased transcription or mRNA instability. The molecular basis of the decrease in *Lyst* mRNA in *bg*²¹ is not yet known, but it is reminiscent of the leaky ablation of mature message associated with an intronic retrotransposition event¹⁴.

The predicted open reading frame (ORF) of *Lyst* was 4,635



sponding to D13Sf1 to D13Sf18, respectively¹³. *Lyst* was isolated from YAC 195A8, a 650-kb clone, by using direct cDNA selection¹³. The physical location of *Lyst*-associated STSs on YAC and P1 clones are shown in red (MGD accession number MGD-PMEX-13).

FIG. 2 Intragenic deletion of *Lyst* in *bg^{11J}* DNA. a–c, Southern blot identification of an intragenic *Lyst* deletion in *bg^{11J}* DNA. A Southern blot was sequentially hybridized¹¹ with 3 *Lyst* probes: a, the probe (nucleotides 1,262–3,433 of *Lyst* cDNA) extends upstream from the *bg^{11J}* deletion; b, the probe (nucleotides 2,835–3,433 of *Lyst* cDNA) is completely deleted; c, the probe (nucleotides 3,594–4,237 of *Lyst* cDNA) extends downstream from the *bg^{11J}* deletion. Restriction endonucleases are indicated at the bottom of each panel, and molecular size standards (in kb) are shown to the left. Similar results were obtained with 3 additional restriction endonucleases (data not shown). The *bg^{11J}* mutation was discovered in 1992 at The Jackson Laboratory in a C57BL/6J-*jb* mouse at generation N4 after transfer from B6C3Fe-a/a. The mutation *jb* had, in turn, been discovered 14 generations earlier in B6C3Fe-a/a-*hyh* mice at generation N3 after transfer from C57BL/10J. The *hyh* mutation arose in C57BL/10J mice, and was maintained in that strain until transfer at F15. Thus the possible contributors of genetic information to *bg^{11J}* include C57BL/6J, C3HeB/FeJ and C57BL/10J. Southern blots were prepared from genomic DNA of all potential progenitor mouse strains, but only C57BL/10J, C57BL/6J and C57BL/6J-*bg^{11J}* are shown. d–f, PCR analysis of the *bg^{11J}* deletion. C57BL/10J, C3HeB/FeJ, C57BL/6J and C57BL/6J-*bg^{11J}* genomic DNA and *Lyst* cDNA were used as templates in the PCR reactions. Amplicons illustrated correspond to: d, *Lyst* cDNA nucleotides 1,337–1,837, which represent exon β and are upstream from the deletion; e, nucleotides 2,670–3,210, which represent exon γ , which is deleted in *bg^{11J}* DNA; and f, nucleotides 4,913–5,433, which represents an exon downstream from the deletion. No amplicon was observed in control PCR reactions performed without template. More than 30 other STSs that had been localized within the *bg* non-recombinant interval PCR amplified normally from *bg^{11J}* DNA. g, Genomic structure of *Lyst* in the vicinity of the *bg^{11J}* deletion. *Lyst* exons (α , β , γ , δ , ϵ and ϕ) are depicted by black boxes, and intervening introns by a solid line. Nucleotides of the mouse *Lyst* cDNA that correspond to exonic boundaries are indicated above the boxes. The 3' end of exon β , and all of exons γ and δ , are deleted in *bg^{11J}* DNA. The region of *Lyst* protein that is deleted in *bg^{11J}* contains a pair of helices with N-terminal phosphorylation sites. Genomic structure and intronic sequences were ascertained by sequence analysis of nested PCR products, performed with exonic primers and P1 clone DNA as template¹⁴. Boundaries of the *bg^{11J}* deletion were determined by PCR of genomic DNA.



nucleotides, encoding a protein of 1,545 amino acids and relative molecular mass 172,500 (M_r 172.5K) (Fig. 3a). Nucleotides 51–74 are rich in CG dinucleotides, a common feature of the 5' region of housekeeping genes. Comparison with DNA databases indicated that *Lyst* is novel, and resembles only uncharacterized human expressed sequence tags (ESTs) (L. Hillier *et al.*, unpublished data; J. Macke *et al.*, unpublished data). The sequence of a cDNA clone corresponding to one such human EST (Genbank accession number L77889) matched the 5' region of mouse *Lyst* (nucleotide identity was 76% in the 5' untranslated region (UTR), 91% in the ORF, and amino-acid identity was 97%; Fig. 3c); another human EST matched the 3' region of the mouse *Lyst* coding domain (Genbank accession number W26957). On hybridization to PFGE Southern blots of mouse DNA¹², the human clones identified restriction fragments that were indistinguishable from mouse *Lyst*; physical mapping of the human clones to the same region of the mouse genome as *Lyst* indicates that they are indeed homologous to *Lyst*.

It has been suggested that CHS and *bg* represent homologous disorders, as their clinical features¹ and defects in lysosomal transport⁴ are identical. Homology of *bg* and CHS is supported by genetic complementation studies: fusion of fibroblasts from *bg* mice and CHS patients fails to reverse lysosomal abnormalities, in contrast to fusions with normal cells¹⁵. Furthermore, recent genetic linkage studies have shown that CHS maps within a linkage group conserved between human Chromosome 1q43 and the *bg* region on mouse Chromosome 13 (F. J. Barrat *et al.*, manuscript submitted). Therefore *LYST* mutations in CHS patients were sought by sequencing *LYST* lymphoblast and fibroblast cDNAs corresponding to these ESTs from 10 CHS

patients. In one patient, a single-base insertional mutation was found at nucleotides 117–118 of the *LYST* coding domain, resulting in a frame shift and termination after amino acid 62 (Fig. 3c).

Previous studies showing spontaneous aggregation of membrane-bound concanavalin A (capping) suggest that there is a defect in microtubule dynamics in *bg* cells^{16,17}. In a search of the SWISSPROT database, using BLITZ and BLASTP, a similarity was found between a domain in *Lyst* and stathmin (oncoprotein 18), a phosphoprotein that may regulate polymerization of microtubules¹⁰ (27% identity from residues 463 to 536; best expected occurrence by chance, 4.36×10^{-6}). The domain in stathmin that matches *Lyst* is helical and has heptad repeats that participate in coiled-coil interactions with other proteins⁸⁻⁹. The stathmin-like region of *Lyst* is also predicted to be helical and form coiled coils. However, it is the charged residues, rather than the hydrophobic ones, that are conserved between *Lyst* and stathmin, suggesting that the sequence similarity is not primarily due to conserved secondary structure. Thus this region of *Lyst* potentially encodes a coiled-coil protein-interaction domain that may regulate microtubule-mediated lysosome transport. Although *Lyst* is not predicted to have transmembrane helices, the C-terminal tetrapeptide (CYSP; amino acids 1,542–1,545) is strikingly similar to known prenylation sites¹⁸, which could provide attachment to lysosomal/late endosomal membranes through thioester linkage with the cysteine.

Previous studies of *bg* leukocytes have shown correction of microtubule function (as assessed by Concanavalin A capping) and natural killer activity when treated with inhibitors of protein kinase C (PKC) breakdown^{19,20}, suggesting that *bg* might be

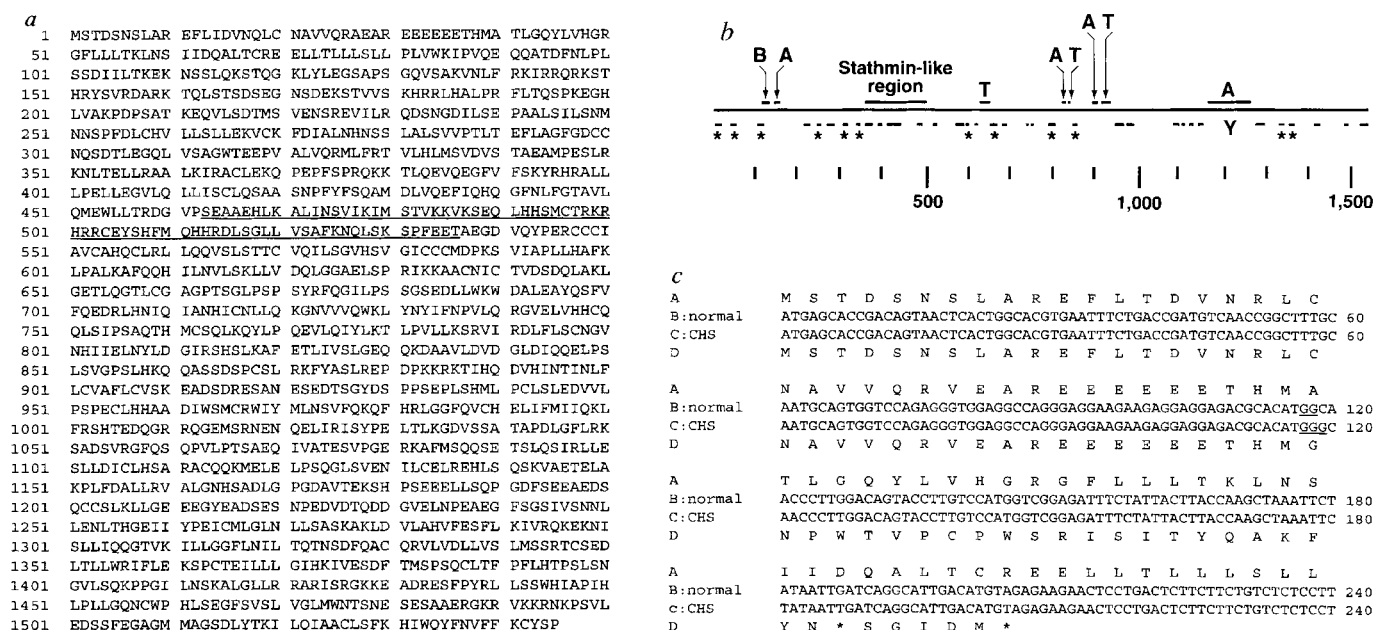
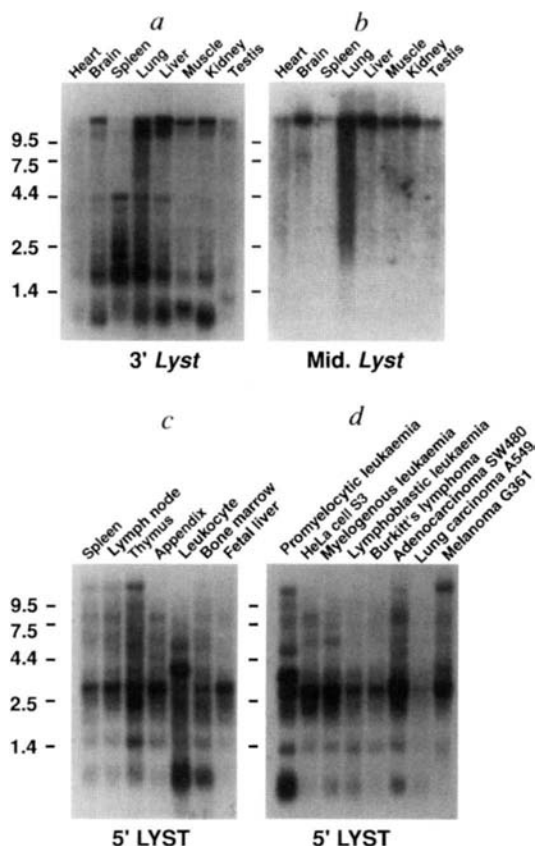


FIG. 3 Amino-acid sequence and predicted peptide structure of mouse and human *Lyst* cDNAs. **a**, Predicted amino-acid sequence of the mouse *Lyst* cDNA. The region (residues 463–536) with sequence similarity to stathmin is underlined, as is a C-terminal prenylation motif. **b**, Predicted secondary structure^{27–30} and motifs of mouse *Lyst* protein. The solid line denotes the protein sequence. Above the line are the stathmin-like region and areas rich in acid (A; Asp, Glu), basic (B; His, Arg, Lys), and turn (T; Pro, Gly, Ser, Thr, Asp) residues. The latter, especially at regions 600–625 and 875–895, are too long to be a simple turn between secondary structural elements, and are either very flexible or very rigid (in a manner reminiscent of collagen, although they do not fit the collagen consensus motifs). Helices predicted

with >80% confidence are shown as solid lines below the protein sequence, and asterisks identify helices with N-terminal phosphorylation motifs. A tyrosine phosphorylation motif (residue 1,214) is designated Y. **c**, Nucleotide and predicted amino-acid sequence of the 5' end of the human *LYST* coding domain showing a frame-shift mutation in a patient with CHS. Line A illustrates the predicted normal human amino-acid sequence; line B shows the normal human nucleotide sequence; line C shows the nucleotide sequence from a patient with CHS; line D shows the predicted amino-acid sequence from the CHS patient. The CHS nucleotide sequence contains an insertional mutation at nucleotide 117–118 that results in a frame shift and premature termination (asterisk).



regulated by phosphorylation. *Lyst* contains 25 sites of potential phosphorylation by PKC, 36 by casein kinase II (CKII) (many of which overlap those of PKC), two by cAMP-dependent protein kinase, and one by tyrosine kinase (Fig. 3b). Almost half of the predicted helices outside the stathmin-like region (14 of 30) have a PKC- or CKII-phosphorylation signal at their amino terminus, and eight of them form consecutive helical pairs. Thus *Lyst* seems to contain helical bundles with clusters of phosphorylation sites at either end. Stathmin also has an N-terminal phosphorylation site and helix motif, and these *Lyst* domains may have a similar 'signal-relaying' function to stathmin^{8,9}. Furthermore, phosphorylation of these positions could provide a control mechanism by causing a conformational shift in the bundles, thereby affecting interactions with other molecules.

Northern analysis and RT-PCR indicated that *Lyst* is ubiquitously transcribed, both temporally and spatially, in mouse and human tissues (Fig. 4). Northern blot analysis also revealed complex alternative splicing of *Lyst* mRNA, with both constitutive and anatomically restricted *Lyst* mRNA isoforms. The largest *Lyst*

FIG. 4 Northern blot analysis of mouse and human *Lyst*. **a**, **b**, Northern blots of 2 µg poly(A)⁺ RNA from various mouse tissues (Clontech) hybridized with probes that correspond to **a**, nucleotides 4,423–4,631, and **b**, nucleotides 1,430–2,457 (exon β) of mouse *Lyst* cDNA. **c**, **d**, Northern blot of 2 µg poly(A)⁺ RNA from **c**, various human lymphoid tissues, or **d**, human cancer cell lines, hybridized with a probe that corresponds to nucleotides 357–800 of human *LYST* cDNA. Molecular size standards (in kb) are shown to the left. Hybridization of mouse mRNA with probes from mouse *Lyst* exons α and γ gave identical results to those shown with exon β (**b**), whereas probes from exons δ, ε, and φ gave results identical to those shown in **a**.

transcript in human and mouse was 12–14 kb, but this transcript was not constitutively expressed. In mRNA from mouse spleen, human peripheral blood leukocytes, promyelocytic leukaemia HL-60, and several leukaemia lines, the 12–14-kb isoform was either undetectable or barely detectable, but smaller *Lyst* transcripts were abundant (Fig. 4). Given the significance for *bg* mice and CHS patients of defects in the lysosomal and late-endosomal compartments of granulocytes, NK cells and cytolytic T lymphocytes^{21–24}, it is likely that these *Lyst* mRNAs of ~3 kb and 4 kb represent the transcripts of primary functional significance.

Probes derived from the 5' or 3' ends of the *Lyst* ORF detected multiple bands on northern blots (Fig. 4a, c), whereas probes that corresponded to more central exons (β and γ) identified only the 12–14-kb transcript (Fig. 4b). Assuming that mature, cytoplasmic *Lyst* transcripts were detected, these results suggest modular assembly of *Lyst* mRNAs, with obligatory usage of the exons that contain start and termination codons, and adaptability in retention of central exons. The possible coexistence of several *Lyst* protein isoforms within cells would allow precise control of functions. □

Received 10 May; accepted 20 June 1996.

1. Blume, R. S. & Wolff, S. M. *Med. Baltimore* **51**, 247–280 (1972).
2. Zhao, H. et al. *Lab. Invest.* **71**, 25–34 (1994).
3. Brandt, E. J., Elliott, R. W. & Swank, R. T. *J. Cell Biol.* **67**, 774–788 (1975).
4. Burkhardt, J. K., Wiebel, F. A., Hester, S. & Argon, Y. *J. exp. Med.* **178**, 1845–1856 (1993).
5. Willingham, M. C., Spicer, S. S. & Vincent, R. A. *Exp. Cell Res.* **136**, 157–168 (1981).
6. Holcombe, R. F., Jones, K. L. & Stewart, R. M. *Immunodeficiency* **5**, 131–140 (1994).
7. Swank, R. T. & Brandt, E. J. *Am. J. Path.* **92**, 755–769 (1978).
8. Sobel, A. *Trends biochem. Sci.* **16**, 301–305 (1991).
9. Maucuer, A., Camonis, J. H. & Sobel, A. *Proc. natn. Acad. Sci. U.S.A.* **92**, 3100–3104 (1995).
10. Belmont, L. D. & Mitchison, T. J. *Cell* **84**, 623–631 (1996).
11. Barbosa, M. D. F. S. et al. *Genomics* **30**, 439–444 (1995).
12. Kingsmore, S. F. et al. *J. invest. Med.* (in the press).
13. Kingsmore, S. F. et al. *Mamm. Genome* (in the press).
14. Kingsmore, S. F. et al. *Nature Genet.* **7**, 136–142 (1994).
15. Perou, C. M. & Kaplan, J. *Somat. Cell molec. Genet.* **19**, 459–468 (1993).
16. Oliver, J. M., Zurier, R. B. & Berlin, R. D. *Nature* **253**, 471–473 (1975).
17. Oliver, J. M. & Zurier, R. B. *J. clin. Invest.* **57**, 1239–1247 (1976).
18. Clarke, S. A. *Rev. Biochem.* **61**, 355–386.

19. Sato, A. et al. *J. Leuk. Biol.* **48**, 377–381 (1990).
20. Ito, M. et al. *Biochem. biophys. Res. Commun.* **160**, 433 (1989).
21. Gallin, J. I., Bujak, J. S., Patten, E. & Sheldon, M. W. *Blood* **43**, 201–206 (1974).
22. Roder, J. & Duwe, A. *Nature* **278**, 451–453 (1979).
23. Saxena, R. K., Saxena, Q. B. & Adler, W. H. *Nature* **295**, 240–241 (1982).
24. Baetz, K., Isaza, S. & Griffiths, M. J. *Immun.* **154**, 6122–6131 (1995).
25. Kusumi, K. et al. *Mamm. Genome* **4**, 391–392 (1993).
26. Pierce, J. C., Sauer, B. & Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2056–2060 (1992).
27. Gibrat, J. F., Garnier, J. & Robson, B. *J. molec. Biol.* **198**, 425–443 (1987).
28. Deleage, G. & Roux, B. *Protein Engng* **1**, 289–294 (1987).
29. Geourjon, C. & Deleage, G. *Protein Engng* **7**, 157–164 (1994).
30. Rost, B. & Sander, C. *Proteins* **19**, 55–72 (1994).

ACKNOWLEDGEMENTS. We thank H. Sweet for investigating the origin of *bg*¹²¹; K. Achey for technical assistance; and S. Segal for inspiration. M.L. is supported by the National Center for Human Genome Research. S.J.B. is supported by the National Institutes of Health and the Vanderbilt Cancer Center. S.F.K. is supported by the American Cancer Society, the Arthritis Foundation, Glaxo-Wellcome Research and Development, and the National Institutes of Health.

CORRESPONDENCE and requests for materials should be addressed to S.F.K. (e-mail: kingsmore@cmg.health.ufl.edu).

ICAM-2 redistributed by ezrin as a target for killer cells

Tuula S. Helander*, Olli Carpén*, Ossi Turunen†, Panu E. Kovanen*, Antti Vaheri† & Tuomo Timonen*‡

Departments of * Pathology, and † Virology, University of Helsinki, Haartman Institute, PO Box 21 (Haartmaninkatu 3), FIN-00014 Helsinki, Finland

VERY little is known about the receptors and target molecules involved in natural killer (NK) cell activity. Here we present a model system in which interleukin-2-activated killing by NK cells depends on the intercellular adhesion molecule ICAM-2 and is regulated by the distribution of ICAM-2. The level of ICAM-2 expression in NK-sensitive and resistant cells is similar, but in sensitive cells ICAM-2 is concentrated into bud-like cellular projections known as uropods, whereas in resistant cells it is evenly distributed. The cytoskeletal-membrane linker protein ezrin is also localized in uropods. Transfection of human ezrin into NK-resistant cells induces uropod formation, redistribution of ICAM-2 and ezrin, and sensitizes target cells to interleukin-2-activated killing. These results reveal a new mechanism of target-cell recognition: cytotoxic cells recognize adhesion molecules that are already present on normal cells, but in diseased cells are concentrated into a biologically active cell-surface region by cytoskeletal reorganization. The results also highlight the importance of cytoskeletal interactions in the regulation of ICAM-2-mediated adhesive phenomena.

We have shown earlier that mouse thymoma BW5147 cells, which are resistant to human interleukin-2 (IL-2)-activated killer cells, become highly susceptible when hybridized with human

chromosome 6 (ref. 1). The parent cells are not bound by the killer cells, whereas 30–40% of IL-2-activated peripheral blood lymphocytes attached to the chromosome-6-hybrid cells. Consequently, chromosome 6 may carry a gene(s) that encodes either target molecules for activated killer cells or elements regulating the function of target molecules.

We first investigated which effector and target molecules were involved in recognition of the 6-hybrids. As shown in Fig. 1, cytotoxicity of IL-2-activated lymphocytes to 6-hybrids was efficiently inhibited by monoclonal antibodies against the α -(CD11a) and β -(CD18) chains of the human β_2 -integrin CD11a/CD18 (present on effector cells), and by antibodies against mouse ICAM-2 (the integrin ligand present on target cells). Function-blocking monoclonal antibodies against the other α -chains (CD11b and CD11c) of β_2 -integrins did not affect the killing, neither did antibodies against mouse ICAM-1. In line with human effector cells, IL-2-activated murine killer cells (isolated from AKR mouse spleen) lysed the 6-hybrids, but did not recognize the parental cells (data not shown). These results indicate that the CD11a/CD18–ICAM-2 pathway could be instrumental in the recognition of 6-hybrids by IL-2-activated killer cells.

Using flow cytometry, the parent cells and 6-hybrids were found to express equal amounts of mouse ICAM-2 (Fig. 2a). In immunofluorescence microscopy, however, mouse ICAM-2 was observed to be evenly distributed on the plasma membrane of the parent cells, whereas the 6-hybrids had a prominent, usually single, uropod (bud-like extension), on the tip of which ICAM-2 was concentrated (Fig. 2b, c). The results show that mouse ICAM-2 was redistributed in BW5147 cells as a result of the introduction of human chromosome 6. When killer cell/6-hybrid cell conjugates were analysed by phase-contrast microscopy, 70% of the conjugates were uropod-associated, showing that the majority of IL-2-activated killer cells recognize target structures expressed on the uropods (Fig. 3a). As the cytotoxicity to the 6-hybrids was inhibited by anti-ICAM-2 antibodies, we conclude that the redistributed mouse ICAM-2 is involved in the recognition of the 6-hybrids by IL-2-activated killer cells.

Association with the cytoskeleton regulates the distribution and

‡ Present address: Imperial College of Science, Technology and Medicine, Wolfson Laboratories, Department of Biochemistry, London SW7 2AY, UK.

function of transmembrane surface structures². To study the mechanism of ICAM-2 redistribution in our model, we screened candidate genes on chromosome 6 that might encode cytoskeletal proteins. The gene encoding ezrin, a membrane-organizing protein proposed to function as a membrane-actin cytoskeleton linker, has been localized on chromosome 6q22-q27 (refs 3,4). Both immunofluorescence microscopy and immunoblotting indicated that 6-hybrids did in fact express human ezrin (Figs 2*d,e*, 3*b,c*). Notably, ezrin colocalized with ICAM-2 in the uropods of 6-hybrids. Disruption of the F-actin cytoskeleton by cytochalasin B resulted in disappearance of uropods and the concentrated distribution of ICAM-2 (data not shown). Thus, cell-surface distribution of ICAM-2 is controlled by the cytoskeleton, with ezrin possibly being involved in the interaction.

To clarify this role of ezrin in the redistribution of ICAM-2 and target-cell sensitization, human ezrin complementary DNA was transfected into parent murine BW5147 cells and into the human embryonal kidney cell line 293. Figure 4 shows that ezrin transiently expressed in BW5147 cells induced uropod formation and redistribution of ICAM-2 in the uropods. Human ezrin in the transfected cells was located in uropods, as seen in the 6-hybrid cells (Fig. 4*d*). The transfected cells (both mouse BW5147 and

human 293 cells) were sensitized to IL-2-activated killer cells to the degree predicted from calculations of transfection efficiency (Fig. 4*e,f*). These results are direct evidence that ezrin is essential in the redistribution of ICAM-2 and sensitization of target cells to IL-2-activated killer cells.

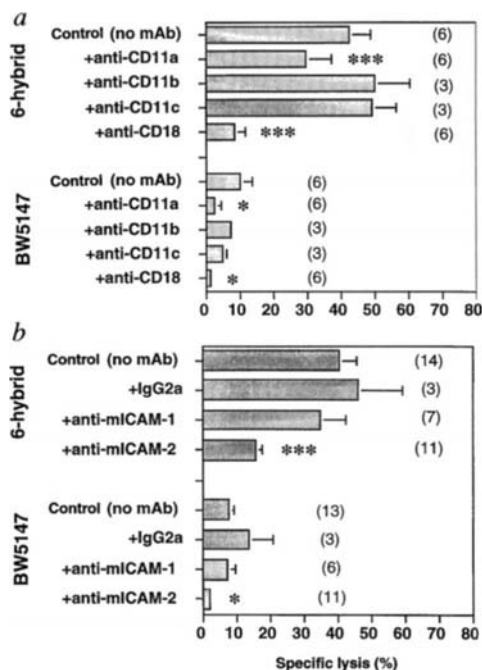


FIG. 1 The sensitivity of 6-hybrids and parent BW5147 cells to lysis by IL-2-activated human peripheral blood lymphocytes (PBL). The effector cells (a) and target cells (b) are preincubated with different monoclonal antibodies (mAbs). The results are expressed as mean values ($\pm$ standard error) and the number of experiments is shown in parentheses. Significance of differences (* $P < 0.05$; *** $P < 0.001$) was determined by the non-parametric Mann-Whitney test.

METHODS. The mouse/human 6-hybrids were generated by fusing human concanavalin A-activated PBL with mouse thymoma BW5147 cells¹. The hybrid populations used contained human chromosome 6 in at least 75% of the cells. The presence of human chromosome 6 was identified by trypsin-Giemsa banding. Human PBL were isolated as described and activated by recombinant IL-2 (500 U ml⁻¹; EuroCetus) overnight²⁹. a, Effector cells were pretreated with 20 μ g ml⁻¹ mAbs to CD11a (TS1/22; from T. Springer), CD11b (60.1; P. Beatty; and LM2/1; T. Springer), CD11c (SHCL-3; R. Schwarting), CD18 (IB4; M. Patarroyo), or no antibody (control). Chromium release was measured after 4 h and assayed at three effector:target cell ratios (50:1 shown here) in triplicate in 96-well round-bottomed microplates. b, The lysis of target cells preincubated with 20 μ g ml⁻¹ of blocking mAbs against mouse ICAM-1 (Serotec; Cymbus Bioscience), mouse ICAM-2 (Pharmingen), control antibody rat IgG2a (Serotec) or no antibody.

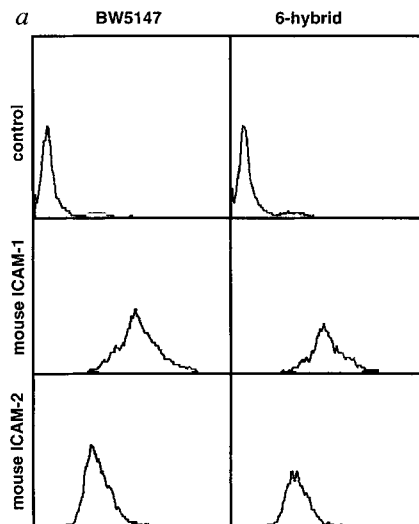


FIG. 2 Expression of mouse ICAMs 1 and 2 (a) and distribution of mouse ICAM-2 (b, c) and ezrin (d, e) on parent cells (b, d) and 6-hybrid (c, e) cells. Two of the uropods in 6-hybrid cells are indicated by arrows. 6-hybrid cells were also stained with mAbs against human ICAMs 1, 2 and 3, but no expression was detected (data not shown).

METHODS. a, Cells were labelled with 10 μ g ml⁻¹ of mAbs for 30 min, and subsequently incubated with FITC-conjugated secondary F(ab')₂ fragment of rabbit anti-rat IgG (Serotec) for 30 min on ice. After washing, fluorescence was analysed by flow cytometry (FACScan, Becton Dickinson). Rat IgG2a served as a control antibody. b and c, Cells were labelled with 10 μ g ml⁻¹ of mouse ICAM-2 antibody for 30 min on ice, and subsequently incubated with FITC-conjugated secondary F(ab')₂ sheep anti-rat IgG2a (Serotec). After washing, cells were cytocentrifuged onto objective slides, fixed for 10 min in 3.5% paraformaldehyde on ice and washed with PBS. d and e, Cells were permeabilized for 5 min in 0.1% Triton-X 100 at +4 °C, and then stained with ezrin antibody (bl. IV; O.T. et al., manuscript in preparation), washed and then incubated with secondary TRITC-conjugated goat anti-rabbit IgG (Cappel). After washing the slides in PBS, a few drops of DABCO (Sigma) were pipetted onto the slides, covered with a coverslip, analysed by fluorescence microscopy and photographed.

Recognition of target cells by NK cells can be divided in two phases, target-cell binding and triggering of cytotoxicity. Surface molecules mediating the recognition have only partially been identified. Binding occurs mostly through the CD11/CD18-ICAM and CD2-LFA-3 pathways. Triggering depends on a balance between inhibitory and stimulatory signals^{5,6}. The principal inhibitory pathways are Ly-49, NKAT and CD94-MHC class

I-peptide complexes⁶⁻¹⁰. Triggering receptors that mediate positive signals are less well characterized. Candidate receptors, such as NKR-P1, NK-TR1, 2B4 and p38 have been identified, but their target molecules have remained elusive^{6,10-16}.

The order of binding and triggering at the start of NK cytotoxicity is not entirely clear. By analogy with T-cell recognition, it is believed that NK cytotoxicity starts by signalling through trigger-

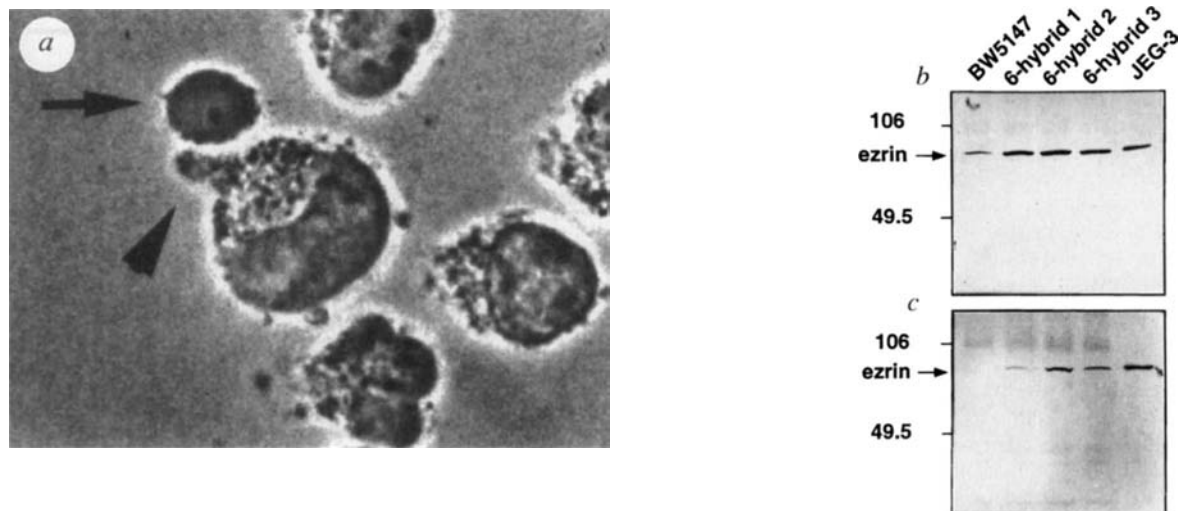


FIG. 3 a, Phase-contrast micrograph of an IL-2-activated lymphocyte (arrow) adhering to the uropod (arrowhead) of a 6-hybrid cell. b, c, immunoblotting of human ezrin with 3C12 (b) and CVP1 (c) antibodies. mAb CVP1 is specific for human ezrin and 3C12 also reacts with mouse ezrin (details of mAbs to be described; O.T. *et al.*, manuscript in preparation). Relative molecular mass markers are indicated on the left (in thousands).

METHODS. a, Effector and target cells were mixed at a 1:1 ratio, incubated for 60 min, cytocentrifuged onto slides, fixed and photographed. b and c, Proteins from parental BW5147 cells, 6-hybrids (three different clones carrying human chromosome 6) and JEG-3 choriocarcinoma cells (positive control) were immunoblotted with mAbs: b, 3C12 (1:20 000) c, and CVP1 (1:5000).

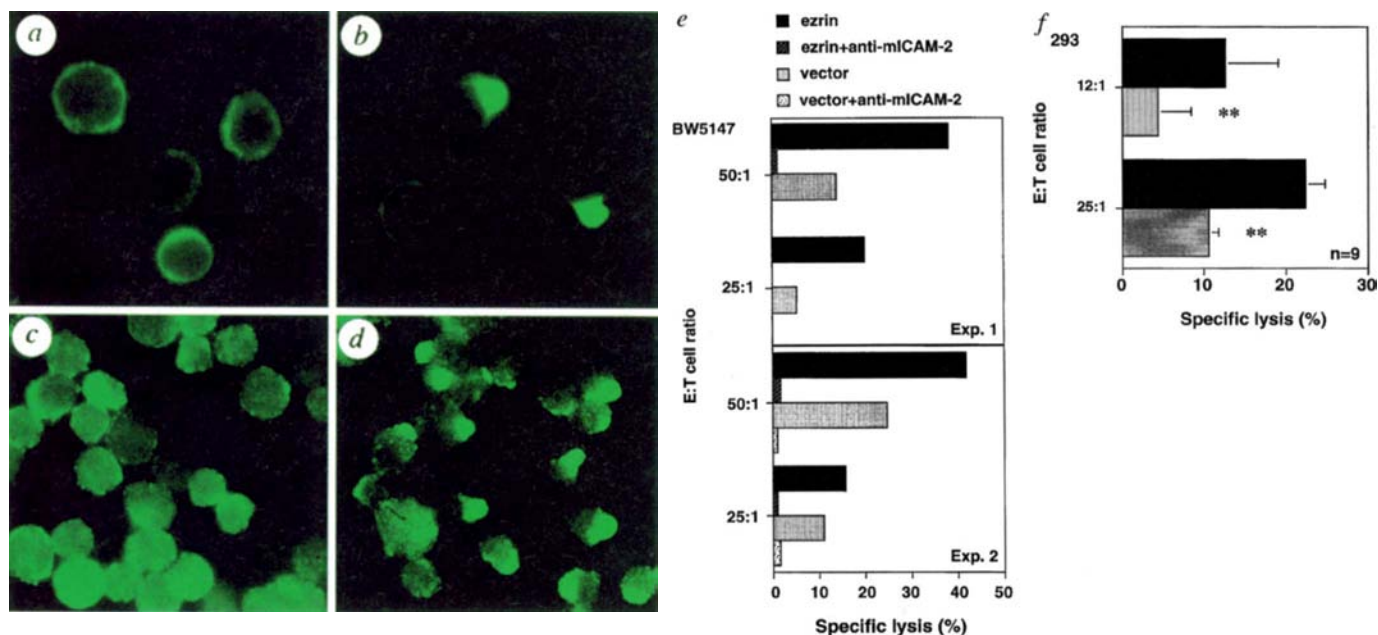


FIG. 4 Distribution of mouse ICAM-2 and ezrin, and target sensitivity of cells transfected with human ezrin cDNA. Note formation of uropods in ezrin transfectants (b, d) in comparison to the round morphology of mock-transfected (a, c) BW5147 cells. Both mouse ICAM-2 (b) and ezrin (d) are invariably concentrated into the uropod region. e, Two cytotoxicity experiments with ezrin cDNA and vector-transfected BW5147 cells blocked by mouse ICAM-2 mAb ($20 \mu\text{g ml}^{-1}$); and f, 9 experiments with transfected 293 cells. The transfection efficiency in all experiments was between 20 and 25%.

and 25%. Significance of differences (** $P < 0.01$) was determined by the non-parametric Mann-Whitney test.

METHODS. Human ezrin cDNA was subcloned into the pcDNA3 eukaryotic expression vector (Invitrogen). Cells were transfected with DOTAP-transfection reagent according to the manufacturer's instructions (Boehringer). Analysis and experiments were carried out 48 h after transfection. Staining for mouse ICAM-2 (a, b) and ezrin (c, d), and determination of cytotoxicity were done as already described.

ing receptors, leading to a conformational change in CD11/CD18. This facilitates binding and enhances triggering. The model implies that triggering is the crucial step in cytotoxicity, binding being secondary, supportive and non-selective.

Our model, however, implies that binding is the selective, rate-limiting step. The prerequisite of this type of recognition is the presence of IL-2. It is probably this cytokine, instead of triggering receptors, that induces the conformational change of CD11a/CD18. Our data do not allow us to distinguish whether triggering of cytotoxicity is transduced through CD11a/CD18, or whether there is a separate triggering pathway that becomes functional only in conjunction with ezrin-induced uropod formation.

The upregulation of binding and cytotoxicity through target-cell ligand redistribution is a new concept in killer-cell biology and perhaps in other ICAM-2-associated adhesion phenomena as well. It may explain why it has been so difficult to find 'foreign' NK target structures: the targets are already present on normal cells, only to become recognizable in abnormal planar distribution.

As specific cell-surface projections, uropods are seen in lymphoid cells, and several cytoskeleton-associated surface molecules are concentrated in these structures¹⁷⁻¹⁹. Our results show that ezrin, a morphogenic membrane-cytoskeleton linker molecule regulates uropod formation. The mechanism of ezrin-controlled uropod formation and ICAM-2 redistribution remains to be clarified. The redistribution of ezrin in 6-hybrids or in transfected cells could be controlled by overexpression, homodimerization, or dimerization with other members of the ERM (ezrin-radixin-moesin) family or by phosphorylation²⁰. Ezrin associates with F-actin present in the cytoskeleton and with the cell-surface proteins CD44 and ICAM-1 (CD54) (refs 21, 22; L. Heiska *et al.*, manuscript in preparation). It links the cell membrane to cytoskeleton and regulates the distribution and function of cell-surface receptors^{4,22-27}. Several cell-surface molecules may be involved in the cytolytic interaction, but ICAM-2 is critical in target sensitization. Ezrin-induced accumulation of ICAM-2 in uropods, associated with increased effector/target conjugation and cytotoxicity, indicates that leukocyte adhesion can be regulated by cytoskeleton-induced redistribution of existing adhesion molecules on cell surface. Viruses, in particular the herpes viruses, sensitize infected cells to NK activity. Moreover, herpes simplex virus infection induces the formation of microvilli and redistribution of ezrin into the newly formed microvilli in cultured fibroblasts²⁸. We suggest that ezrin redistributes adhesion molecules within the microvilli and other cell-surface projections in virus-infected or otherwise altered cells, and could thus make the local adhesion molecule concentration sufficient for recognition by killer cells. □

Received 16 February; accepted 24 May 1996.

- Helander, T., Timonen, T., Kalliomäki, P. & Schröder, J. *J. Immunol.* **147**, 2063–2067 (1991).
- Arpin, M., Algrain, M. & Louvard, D. *Curr. Opin. Cell Biol.* **6**, 136–141 (1994).
- Turunen, O. *et al. J. Biol. Chem.* **264**, 16727–16732 (1989).
- Algrain, M., Turunen, O., Vaheri, A., Louvard, D. & Arpin, M. *J. Cell Biol.* **120**, 129–139 (1993).
- Kärre, K. *Science* **267**, 978–979 (1995).
- Chambers, W. H. & Brissette-Storkus, C. S. *Chem. Biol.* **2**, 429–435 (1995).
- Trinchieri, G. *J. exp. Med.* **180**, 417–421 (1994).
- Malnati, M. S. *et al. Science* **267**, 1016–1018 (1995).
- Colonna, M. & Samaridis, J. *Science* **268**, 405–408 (1995).
- Gumperz, J. E. & Parham, P. *Nature* **378**, 245–248 (1995).
- Giorda, R. *et al. Science* **249**, 1298–1300 (1990).
- Yokoyama, W. M. & Seaman, W. E. *A. Rev. Immunol.* **11**, 613–635 (1993).
- Frey, J. L. *et al. J. exp. Med.* **174**, 1527–1536 (1991).
- Giordina, S. L. *et al. J. Immunol.* **154**, 80–87 (1995).
- Garni-Wagner, B. A., Purohit, A., Mathew, P. A., Bennet, M. & Kumar, V. *J. Immunol.* **151**, 60–70 (1993).
- Valiante, N. M. & Trinchieri, G. *J. exp. Med.* **178**, 1397–1406 (1993).
- Dougherty, G. J., Murdoch, S. & Hogg, N. *Eur. J. Immunol.* **18**, 35–39 (1988).
- Dustin, M. L., Carpen, O. & Springer, T. A. *J. Immunol.* **148**, 2654–2663 (1992).
- Campanero, M. R. *et al. J. Cell Biol.* **123**, 1007–1016 (1993).
- Bretscher, A. *J. Cell Biol.* **108**, 921–930 (1989).
- Turunen, O., Wahlström, T. & Vaheri, A. *J. Cell Biol.* **126**, 1445–1453 (1994).
- Tsukita, S. *et al. J. Cell Biol.* **126**, 391–401 (1994).
- Takeuchi, K. *et al. J. Cell Biol.* **125**, 1371–1384 (1994).
- Hanzel, D., Reggio, H., Bretscher, A., Forte, J. G. & Mangeat, P. *EMBO J.* **10**, 2363–2373 (1991).
- Gary, R. & Bretscher, A. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10846–10850 (1993).
- Andréoli, C., Martin, M., Le Borgne, R., Reggio, H. & Mangeat, P. *J. Cell Sci.* **107**, 2509–2521 (1994).
- Martin, M. *et al. J. Cell Biol.* **128**, 1081–1093 (1995).

- Pakkanen, R., Bonsdorff, C.-H., Turunen, O., Wahlström, T. & Vaheri, A. *Eur. J. Cell Biol.* **46**, 435–443 (1988).
- Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immunol.* **3**, 645–649 (1973).

ACKNOWLEDGEMENTS. We thank M. Patarroyo for mAbs against mouse ICAM-1 (YN1/2) and human CD18 (IB4); and Jääskeläinen and M. Sainio for discussion and for comments on the manuscript. Supported by the Academy of Finland, Finnish Cancer Society, the Finnish Cultural Foundation, the National Cancer Institute of Finland and the University of Helsinki.

CORRESPONDENCE and requests for materials should be addressed to T.S.H. (e-mail: Tuula.Helander@Helsinki.Fi).

Phosphorylation-dependent activation of the Ras-GRF/CDC25^{Mm} exchange factor by muscarinic receptors and G-protein $\beta\gamma$ subunits

Raymond R. Mattingly & Ian G. Macara

Departments of Pathology, and Microbiology and Molecular Genetics, University of Vermont College of Medicine, and the Vermont Comprehensive Cancer Center, Burlington, Vermont 05405, USA

MUSCARINIC receptors activate Ras through a pathway distinct^{1,2} from that mediated through translocation of the exchange factor mSos1 by receptor tyrosine kinases^{3,4}. Here we report that muscarinic receptors can activate another Ras exchange factor, CDC25^{Mm}, or p140^{Ras-GRF} (refs 5,6). In NIH-3T3 cells expressing subtype 1 human muscarinic receptors (hm1), the agonist carbachol selectively increased the specific activity and phosphorylation state of epitope-tagged Ras-GRF. This stimulation was reversed by protein phosphatase 1 (PP1), and prevented by transducin α -subunits. Carbachol treatment of neonatal rat brain explants increased Ras exchange factor activity and the phosphorylation state of endogenous Ras-GRF. In COS-7 cells, cotransfection of hm1 or hm2 receptors with Ras-GRF conferred carbachol-dependent increases in exchange-factor activity, whereas cotransfection with G-protein $\beta\gamma$ subunits caused a constitutive activation that was sensitive to PP1. These results demonstrate a G-protein-coupled mechanism for Ras activation, mediated by p140^{Ras-GRF}.

Ras proteins function as molecular switches that are activated by exchange of GTP for GDP⁷. A complete pathway for the activation of Ras by receptors with intrinsic tyrosine kinase activity has been defined, involving association of the autophosphorylated receptor with a cytosolic complex of Grb2 and the mSos exchange factor, translocating mSos to the plasma membrane, where it can stimulate nucleotide exchange on Ras^{3,4}. Activation of Ras by G-protein-coupled receptors is less clearly defined. Lysophosphatidic acid activates Ras by means of a pathway sensitive to pertussis toxin and genistein⁸, a tyrosine kinase inhibitor. The mechanism may involve stimulation of a tyrosine kinase by G-protein $\beta\gamma$ subunits and increased mSos association with Shc and Grb2 coupling proteins⁹. Muscarinic receptors, however, use a different pathway. Ras activation by hm2 receptors expressed in Rat-1 cells is resistant to genistein¹, and, in NIH-3T3 cells that stably express hm1 receptors (3T3/hm1), carbachol activates Ras through a pathway distinct from that used by EGF². This hm1-dependent Ras activation does not require inhibition of Ras-GAP activity¹⁰, or mSos translocation, and is not accompanied by increased Shc phosphorylation (data not shown).

The Ras-GRF exchange factor has not been implicated in tyrosine kinase-dependent pathways. Rather, it may be an intermediary between calcium signals and Ras¹¹. Ras-GRF is expressed at high levels in the brain^{5,6,12}, but is also found in other tissues and

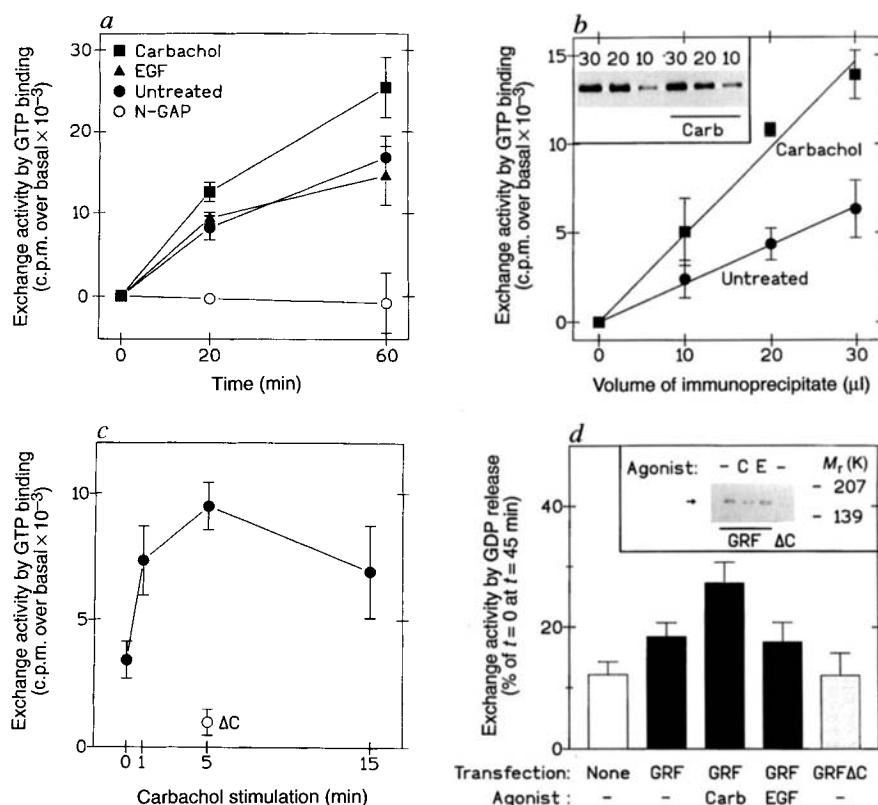
FIG. 1 Carbachol activates Ras-GRF in 3T3/hm1 cells. Carbachol (Carb) treatments were 100 μ M for 5 min or as shown. Data are mean $\pm$ s.e.m. from at least 3 independent transfections, each assayed in triplicate. *a-c*, Stimulated exchange of [α - 32 P]GTP onto c-Ha-Ras by 12CA5 anti-HA1 immunoprecipitates of cells transfected² with 20 μ g per plate pKH3Ras-GRF. The basal rate of exchange, defined by incubation with unreacted beads, was subtracted from the values shown. Transfections with the triple-tagged, non-catalytic regions of Ras-GAP (N-GAP)² or Ras-GRF (Δ C) yield no exchange activity (open circles). *d*, Release of [α - 32 P]GDP from Ras by 12CA5 immunoprecipitates of cells transfected as shown. Insets, representative 12CA5 western blots² of the immunoprecipitates.

METHODS. Plasmids encoding either Ras-GRF or Ras-GRF Δ C, with or without a triple HA1 tag fused at their N termini, were constructed by subcloning the CDC25^{MM}IV sequence¹⁴ into the *Bam*HI/*Eco*RI sites of pKH3 and pRK7 (ref. 2). After transfection², the cells were serum starved overnight, stimulated, rinsed with ice-cold phosphate-buffered saline (PBS), and extracted with 0.5 ml of either: *a*, RIPA (150 mM NaCl, 50 mM Tris, pH 8.0, 1.0% NP-40, 0.5% Na deoxycholate, 0.1% SDS, 1 mM Na₃VO₄, 100 μ M phenylmethylsulphonyl fluoride, 25 μ g ml⁻¹ aprotinin, 25 μ g ml⁻¹ leupeptin); or *b-d*, RIPA plus 50 mM NaF, 50 mM Na β -glycerophosphate, and 20 mM Na₂P₂O₇.

Extracts were centrifuged and incubated sequentially with 2 μ g 12CA5 or mouse IgG and 15 μ l protein A-Sepharose. Immunoprecipitates were washed twice in extraction buffer and twice in reaction buffer (5 mM MgCl₂, 5 mM KCl, 1 mM EGTA, 10 mM HEPES-NaOH, pH 7.4), followed by immunoblotting and assay. Assays were performed at 30 °C in 70 μ l of reaction buffer plus 2 mM dithiothreitol, 0.2 mg ml⁻¹ BSA, 1 mM

some non-neuronal cell lines¹², and can enhance serum-dependent activation of Ras^{13,14} when it is expressed in NIH-3T3 cells. Serum does not, however, increase the specific activity of Ras-GRF measured *in vitro*^{13,14}. We therefore tested whether muscarinic receptors could stimulate Ras-GRF. Ras exchange activity was assayed in immunoprecipitates from 3T3/hm1 cells transiently transfected with a vector that encodes Ras-GRF tagged with a triple haemagglutinin peptide (HA1₃-Ras-GRF). Stimulation of the cells with carbachol increased the Ras exchange activity present in anti-HA1 immunoprecipitates from cell lysates by 50% (Fig. 1*a*; $P < 0.05$). When protein phosphatase inhibitors were present during cell lysis and processing, carbachol induced a rapid and transient two- to threefold increase in Ras-GRF activity, assayed as increased binding of GTP to recombinant c-Ha-Ras (Fig. 1*b*, *c*) or as increased release of GDP from Ras (Fig. 1*d*). EGF treatment did not stimulate the activity of Ras-GRF (Fig. 1*a*, *d*). There was no effect of carbachol treatment on the exchange activity of an immunoprecipitated HA1₃-mSos1 construct³ (data not shown). Acute agonist treatments did not alter the amount of HA1₃-Ras-GRF recovered by immunoprecipitation (Fig. 1 insets). Taken together these results demonstrate that muscarinic receptor activation can significantly increase the specific activity of Ras-GRF.

To determine whether expression of Ras-GRF can potentiate Ras-dependent responses within the cell, carbachol stimulation of mitogen-activated protein kinase (MAPK) activity was assayed. Cotransfection with the dominant-negative N17 mutant of Ras inhibited the ability of carbachol to stimulate a Myc-tagged MAPK construct (Fig. 2*a*). This result differs from that reported



ATP, and either 1 mM unlabelled GTP (for GDP release) or 100 nM unlabelled GTP plus 1 μ Ci [α - 32 P]GTP (3,000 Ci mmol⁻¹) (for GTP binding), and initiated by addition of 1 μ g c-Ha-Ras¹⁶ prebound to either [α - 32 P]GDP (1,000 Ci mmol⁻¹) or unlabelled GDP, respectively. Portions, withdrawn at the times indicated, were terminated by binding to nitrocellulose filters¹⁶.

for hm1-transfected COS cells, in which carbachol-dependent MAPK activation is independent of Ras¹⁵. Expression of Ras-GRF increased both the efficacy and potency of MAPK activation by carbachol (Fig. 2*b*), and the effect of Ras-GRF on MAPK activation was inhibited by the dominant-negative A17 Ras mutant (Fig. 2*c*). (An inhibition was also seen for N17-Ras, but this mutant also reduced the expression of Ras-GRF.)

Serum stimulation of NIH-3T3 cells expressing Ras-GRF induces a hyperphosphorylation of the protein that is blocked by pertussis toxin¹³. We therefore examined the effect of hm1 activation on Ras-GRF phosphorylation, and found that carbachol, but not EGF, increased the phosphorylation state of HA1₃-Ras-GRF (Fig. 3*a*). Treatment of the immunoprecipitated HA1₃-Ras-GRF with PP1 significantly reversed the carbachol-dependent increase in exchange-factor activity (Fig. 3*b*).

Inclusion of phosphatase inhibitors potentiates Ras exchange-factor activity in rat brain cytosol¹⁶. Explant cultures of neonatal rat brains were therefore prepared to assay the effects of muscarinic receptor stimulation on exchange-factor activity and the phosphorylation state of endogenous Ras-GRF. Carbachol induced a significant increase in Ras exchange activity that was enhanced by inhibition of protein phosphatases (Fig. 3*c*), and also increased the phosphorylation state of Ras-GRF immunoprecipitated from explants labelled with [32 P]orthophosphate (Fig. 3*d*). The increases observed are striking because only a fraction of the cells in the culture that express Ras-GRF would also be expected to express muscarinic receptors, and Ras-GRF may not be the only exchange factor present.

Muscarinic receptor subtypes 1, 3 and 5 couple to the G protein

FIG. 2 Ras-GRF can potentiate carbachol-induced MAPK activation in 3T3/hm1 cells. Activation of MAPK activity by carbachol in cells cotransfected² with a plasmid encoding Myc-tagged p42 MAPK (ERK2)²⁶ plus the indicated constructs. **a**, Cotransfections with the dominant-negative mutant of Ras (N17), or, as a negative control, an equivalent mutant of Rab3A (N36). Top, autoradiogram of MAPK activity of 9E10 immunoprecipitates by myelin basic protein (MBP) phosphorylation²⁶. Bottom, expression of the MAPK construct by 9E10 immunoblot²⁶. **b**, Cotransfections with either pRK7Ras-GRF or pRK7Ras-GRFΔC. Data shown are mean $\pm$ s.e.m. from at least 3 independent transfections, and are relative to control, which is defined as activity present in immunoprecipitates from unstimulated cells cotransfected with Ras-GRFΔC. Sigmoid curves were fitted to the data²⁷, demonstrating that the EC₅₀ for carbachol decreased from 4 to 0.4 μ M with Ras-GRF cotransfection. **c**, Cotransfection with both Ras-GRF and dominant-negative Ras and Rab3A mutants. Top, autoradiogram of MAPK activity. Bottom, Ras-GRF expression by 12CA5 immunoblot of cell lysates that had been precipitated with trichloroacetic acid (TCA).

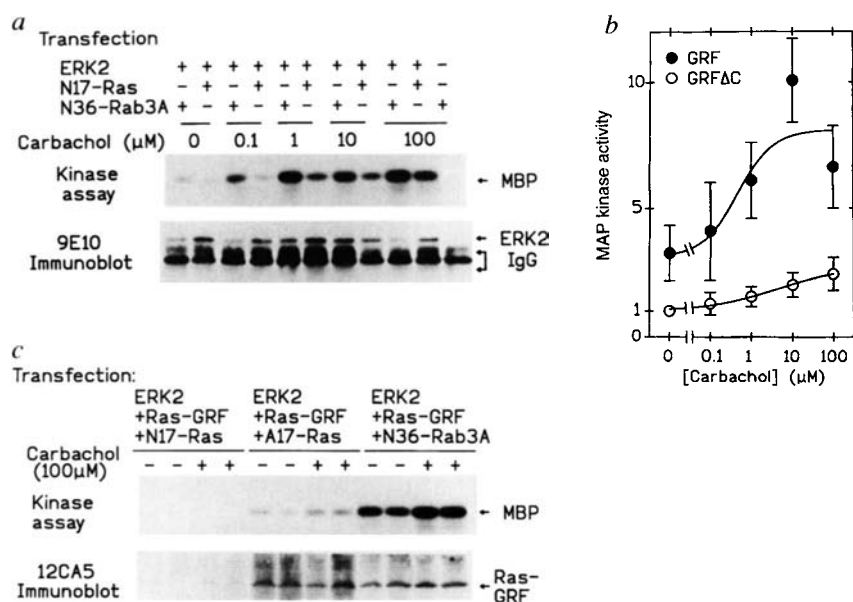
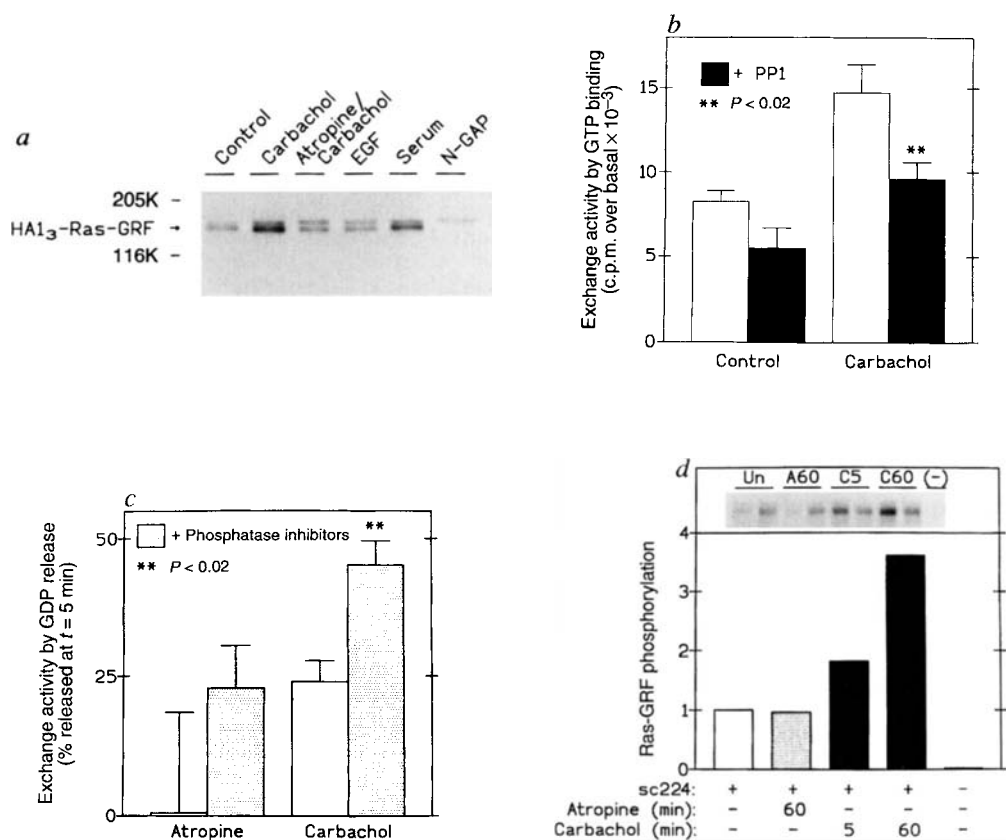


FIG. 3 Carbachol-induced phosphorylation of Ras-GRF correlates with increased exchange activity. **a**, Autoradiogram of HA1₃-Ras-GRF immunoprecipitated from 3T3/hm1 cells labelled² with [³²P]orthophosphate and stimulated as shown. N-GAP shows the result from cells transfected with pKH3N-GAP but without pKH3Ras-GRF (note the nonspecific labelled protein just above the HA1₃-Ras-GRF). Data are representative of 3–7 independent experiments. **b**, PP1 reverses the carbachol-dependent increase in Ras-GRF activity. Data are from 4 independent transfections. **c**, Carbachol-dependent activation of Ras exchange-factor activity in explants of neonatal rat brain. Treatments of 1 μ M atropine and 100 μ M carbachol were for 30 min. Data are from triplicate assays of duplicate cultures. Phosphatase inhibitors were 1 μ M okadaic acid, 1 μ M microcystin, and 1.5 mM *p*-nitrophenyl phosphate. **d**, Carbachol-induced phosphorylation of Ras-GRF in explants of neonatal rat brain. Inset, autoradiogram of sc224 (Santa Cruz) anti-Ras-GRF¹² immunoprecipitates from duplicate cultures treated as shown. Graph shows quantification of the phosphorylation relative to control.



METHODS. **a**, Immunoprecipitates from ³²P-labelled cells were washed 5 times with extraction buffer and once each with PBS and reaction buffer. **b**, Immunoprecipitates, treated with or without 1 unit recombinant PP1 for 30 min at 30 °C in the buffer supplied (New England Biolabs), were assayed for exchange-factor activity. PP1 treatment reduced the phosphorylation state of Ras-GRF that had been immunoprecipitated from ³²P-labelled cells treated with carbachol (data not shown). The PP1 was free of protease activity and did not degrade HA1₃-Ras-GRF (not shown). **c**, **d**, Explant cultures of whole neonatal rat brains were established²⁸ in 24-well plates (1 dispersed brain per well). **c**, Cytosol was prepared as described¹⁶, except that explants were homogenized by passage through a 26-gauge needle and centrifuged at 14,000g, and then

assayed for Ras exchange activity¹⁶. **d**, Explants were labelled for 60 min with 1 mCi ml⁻¹ [³²P]orthophosphate in phosphate-free medium. Ras-GRF was immunoprecipitated with 2 μ g sc224 as described for Fig. 1, except that the RIPA was further supplemented with 1 μ M okadaic acid, 1 μ M microcystin and 1.5 mM *p*-nitrophenyl phosphate. Phosphorylation was quantified by a Bio-Rad GS250 phosphorimager, with the negative control immunoprecipitate defined as background. The values for the Ras-GRF immunoprecipitates were corrected to the total TCA-precipitable radioactivity present in duplicate 1- μ l samples of the lysates.

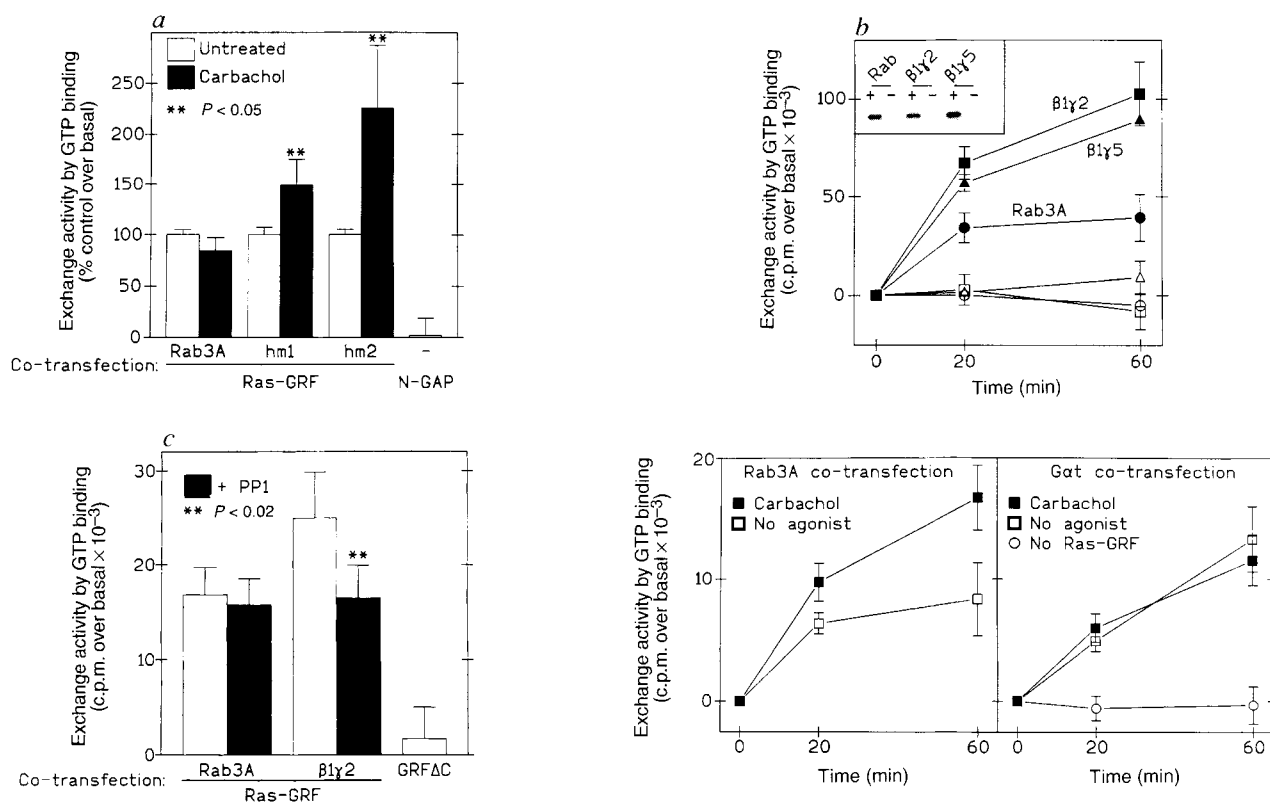


FIG. 4 a, Activation of Ras-GRF by coexpressed hm1 or hm2 in COS-7 cells. Data are from 4 independent transfections. The mean activity present in immunoprecipitates from unstimulated cells was defined as 100%. b, Activation of Ras-GRF by G-protein $\beta\gamma$ subunits coexpressed in COS-7 cells. Filled symbols are 12CA5 immunoprecipitates; open symbols are control (mouse IgG) immunoprecipitates. Data are representative of 4 independent transfections. Inset, western blot with 12CA5 monoclonal antibody² of control (–) or 12CA5 (+) immunoprecipitates of cells expressing HA1₃-Ras-GRF with cotransfection as shown. c, PP1 reverses the Ras-GRF activation induced by $\beta_1\gamma_2$. Data are from duplicate transfections. d, Transducin inhibits carbachol-stimulated Ras-GRF activation in 3T3/hm1 cells. Data are from 3 independent transfections. Carbachol induced a significant ($P < 0.01$ by one-tailed t-test) increase in exchange activity only in the Rab3A cotransfections.

METHODS. Sequences encoding G-protein subunits²⁹ were subcloned into pRK5/pRK7 (gift from D. Pritchett). For β_1 , an *EcoRI/BamHI* fragment was inserted into pRK5. For γ_2 , an *EcoRI/XbaI* fragment was inserted into pRK7. For γ_5 , an *EcoRI* fragment was inserted into pRK7, and the orientation was confirmed by *HindIII* digestion. Overexpression of β_1 was demonstrated in COS-7 cells cotransfected with pRK5 β_1 and either pRK7 γ_2 or pRK7 γ_5 by immunoblotting with MS-1 antiserum (DuPont). COS-7 cells were cotransfected with 2.5 μ g per plate each of pKH3Ras-GRF plus: a, pCDhm1 (ref. 16) or pCDhm2 (ref. 16) or the control pRK7Rab3A(V55E) construct³⁰; or b, c, plus 1.25 μ g per plate each pRK5 β_1 and pRK7 γ_2 or pRK5 β_1 and pRK7 γ_5 subunits. d, 3T3/hm1 cells were cotransfected with 15 μ g per plate each of pKH3Ras-GRF or pKH3N-GAP plus either the control pRK7Rab3A(V55E) construct³⁰ or pCDNA1G α t (gift from E. R. Weiss).

G_q and stimulate phosphatidylinositol turnover and generation of calcium and diacylglycerol signals¹⁷. To test whether these signals could explain the hm1-dependent activation and phosphorylation of Ras-GRF, we investigated whether activation of protein kinase C by 12-*O*-tetradecanoylphorbol-13-acetate (TPA), or increases in cytosolic calcium induced by either ionomycin or thapsigargin, could mimic the effects of carbachol. These treatments failed to increase the activity or the phosphorylation state of HA1₃-Ras-GRF in 3T3/hm1 cells (data not shown). The lack of effect of ionomycin on the specific activity of Ras-GRF agrees with a previous report¹¹.

Expression of HA1₃-Ras-GRF together with hm1 or hm2 receptors in COS-7 cells demonstrates that both of these receptors could confer carbachol-dependent increases in Ras exchange activity (Fig. 4a). Because hm2 receptors couple to G_i rather than G_q^{15,17}, a common element in their signalling could be the G-protein $\beta\gamma$ subunits released from the α subunits on receptor activation. G-protein $\beta\gamma$ subunits are implicated in pathways of Ras and MAPK activation^{9,15,18}, and the overexpression of $\beta_1\gamma_2$ in COS-7 cells is sufficient to activate MAPK^{15,19,20}. Cotransfection of either $\beta_1\gamma_2$ or $\beta_1\gamma_5$ subunits with Ras-GRF induced a marked increase in Ras exchange factor activity ($P < 0.02$), without any increase in HA1₃-Ras-GRF expression (Fig. 4b). Dephos-

phorylation of the immunoprecipitated Ras-GRF by treatment with PP1 reversed the activation induced by $\beta_1\gamma_2$ cotransfection (Fig. 4c).

One method used to implicate $\beta\gamma$ subunits in a signalling pathway is to test whether expression of the transducin α subunit (G α t) can inhibit the response, presumably by sequestering $\beta\gamma$ subunits³⁰. We found that coexpression of G α t with HA1₃-Ras-GRF in 3T3/hm1 fibroblasts abolished the carbachol-induced increase in exchange-factor activity (Fig. 4d).

These results identify Ras-GRF as a potential intermediary between G-protein $\beta\gamma$ subunits generated by muscarinic receptors and Ras. Furthermore, the increase in the specific activity of Ras-GRF is associated with an increase in its phosphorylation state. This pathway to Ras activation is thus distinct from the mSos translocation model for tyrosine kinase-dependent signalling, in which increased phosphorylation of mSos has been associated with downregulation rather than activation²¹. Ras-GRF binds calmodulin and is implicated in a calcium-dependent pathway for Ras activation in cultured cortical neurons treated with ionomycin¹¹. However, neither ionomycin treatment nor calmodulin-binding stimulate *in vitro* exchange activity of Ras-GRF¹¹. Our results therefore suggest that there are at least two mechanisms for Ras activation that converge on the Ras-GRF exchange

factor. Activation of Ras by G-protein-coupled receptors may be of particular relevance in tissues, such as cardiac muscle and the central nervous system^{11,22}, in which Ras activation may be associated with gene expression induced in the hypertrophic growth of muscle or by neurotransmission^{23–25}. □

Received 3 April; accepted 10 June 1996.

- Winitz, S. *et al.* *J. Biol. Chem.* **268**, 19196–19199 (1993).
- Mattingly, R. R., Sorisky, A., Brann, M. R. & Macara, I. G. *Molec. cell. Biol.* **14**, 7943–7952 (1994).
- Egan, S. E. *et al.* *Nature* **363**, 45–51 (1993).
- Buday, L. & Downward, J. *Cell* **73**, 611–620 (1993).
- Martegani, E. *et al.* *EMBO J.* **11**, 2151–2157 (1992).
- Shou, C., Farnsworth, C. L., Neel, B. G. & Feig, L. A. *Nature* **358**, 351–354 (1992).
- Boguski, M. S. & McCormick, F. *Nature* **366**, 643–654 (1993).
- van Corven, E. J., Hordijk, P. L., Medema, R. H., Bos, J. L. & Moolenaar, W. H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1257–1261 (1993).
- van Biesen, T. *et al.* *Nature* **376**, 781–784 (1995).
- Downward, J., Graves, J. D., Wame, P. H., Rayter, S. & Cantrell, D. A. *Nature* **346**, 719–723 (1990).
- Farnsworth, C. L. *et al.* *Nature* **376**, 524–527 (1995).
- Guerrero, I. *et al.* *Oncogene* **12**, 1097–1107 (1996).
- Shou, C., Wurmser, A., Ling, K., Barbacid, M. & Feig, L. A. *Oncogene* **10**, 1887–1893 (1995).
- Cen, H., Papageorge, A. G., Vass, W. C., Zhang, K. & Lowy, D. R. *Molec. cell. Biol.* **13**, 7718–7724 (1993).
- Hawes, B. E., van Biesen, T., Koch, W. J., Luttrell, L. M. & Lefkowitz, R. J. *J. Biol. Chem.* **270**, 17148–17153 (1995).

- Wolfman, A. & Macara, I. G. *Science* **248**, 67–69 (1990).
- Gutkind, J. S., Novotny, E. A., Brann, M. R. & Robbins, K. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4703–4707 (1991).
- Koch, W. J., Hawes, B. E., Allen, L. F. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **91**, 12706–12710 (1994).
- Crespo, P., Xu, N., Simmonds, W. F. & Gutkind, J. S. *Nature* **369**, 418–420 (1994).
- Faure, M., Vaynsenetskaya, T. A. & Bourne, H. R. *J. Biol. Chem.* **269**, 7851–7854 (1994).
- Porfiri, E. & McCormick, F. *J. Biol. Chem.* **271**, 5871–5877 (1996).
- LaMore, V. J. *et al.* *J. Biol. Chem.* **269**, 13490–13496 (1994).
- Hanley, M. R. *Nature* **340**, 97 (1989).
- Ramirez, M. T., Post, G. R., Sulakhe, P. V. & Brown, J. H. *J. Biol. Chem.* **270**, 8446–8451 (1995).
- Thorburn, A. *et al.* *J. Biol. Chem.* **268**, 2244–2249 (1993).
- Posada, J. & Cooper, J. A. *Science* **255**, 212–215 (1992).
- Mattingly, R. R., Wasilenko, W. J., Woodring, P. J. & Garrison, J. C. *J. Biol. Chem.* **267**, 7470–7477 (1992).
- Linnemann, D. in *A Dissection and Tissue Culture Manual of the Nervous System* (ed. Shahar, A. *et al.*) 75–76 (Alan R. Liss, New York, 1989).
- Graber, S. G., Figler, R. A., Kalman-Maltese, V. K., Robishaw, J. D. & Garrison, J. C. *J. Biol. Chem.* **267**, 13123–13126 (1992).
- Brondyk, W. H., McKiernan, C. J., Burstein, E. S. & Macara, I. G. *J. Biol. Chem.* **268**, 9410–9415 (1993).

ACKNOWLEDGEMENTS. We thank M. R. Brann, M. Czech, S. G. Graber, D. R. Lowy, V. May, J. Posada, D. Pritchett, E. R. Weiss and A. Wolfman for reagents; S. A. Callaghan, P. B. Joel, R. R. Perlungher and N. Tatsis for technical advice and assistance; and J. Downward for a critical reading of the manuscript. These studies were supported by the NCI through grants to I.G.M. and a fellowship to R.R.M.

CORRESPONDENCE and requests for materials should be addressed to R.R.M. (e-mail: rmatting@moose.uvm.edu).

Role for c-Abl tyrosine kinase in growth arrest response to DNA damage

Zhi-Min Yuan, Yinyin Huang, Young Whang*, Charles Sawyers*†, Ralph Weichselbaum‡, Surender Kharbanda & Donald Kufe

Division of Cancer Pharmacology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

† Molecular Biology Institute, * Department of Medicine, University of California, Los Angeles, California 90095, USA

‡ Department of Radiation and Cellular Oncology, University of Chicago, Chicago, Illinois 60637, USA

THE c-Abl protein tyrosine kinase is activated by certain DNA-damaging agents¹, and its overexpression causes arrest in the G1 phase of the cell cycle by a mechanism dependent on the tumour-suppressor protein p53 (refs 2–4). Here we investigate the possible role of c-Abl in growth arrest induced by DNA damage. Transient transfection experiments using wild-type or inactivated c-Abl show that both induce expression of p21, an effector of p53, but only wild-type c-Abl downregulates the activity of the cyclin-dependent kinase Cdk2 and causes growth arrest. Exposure to ionizing radiation of cells that stably express active or inactive c-Abl is associated with induction of c-Abl/p53 complexes and p21 expression. However, cells expressing the dominant-negative c-Abl mutant and cells lacking the c-abl gene are impaired in their ability to downregulate Cdk2 or undergo G1 arrest in response to ionizing radiation. We also show that expression of c-Abl kinase in *p21*^{-/-}, but not in *p53*^{-/-}, cells results in downregulation of Cdk2. Our results suggest that c-Abl kinase contributes to the regulation of growth arrest induced by ionizing radiation by a p53-dependent, p21-independent mechanism.

To determine whether c-Abl affects the function of p53, we co-transfected MCF-7 cells with a construct containing the luciferase gene driven by a p53 enhancer from the *MDM2* promoter⁵, and vectors expressing wild-type c-Abl, a kinase-inactive K(290)R mutant² or a kinase-active mutant, designated ΔProl, which is deleted at the p53-binding domain⁴. Co-transfections of the reporter with wild-type or kinase-inactive c-Abl(K-R), but not c-AblΔProl, resulted in induction of luciferase activity (Fig. 1a). By contrast, c-Abl expression had no detectable effect on activa-

tion of an SV40-luciferase construct (Fig. 1a). As transcription of *p21* is regulated by p53 (ref. 6), we investigated whether transfection of c-Abl induces p21. We found that p21 expression increased in the c-Abl and c-Abl(K-R), but not the c-AblΔProl, transfectants (Fig. 1b). Despite induction of p21, which is an inhibitor of Cdk2

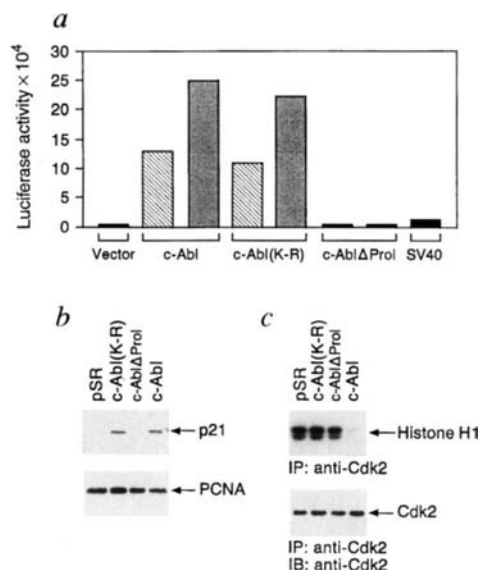


FIG. 1 Overexpression of kinase active c-Abl downregulates Cdk activity. **a**, MCF-7 cells were transfected with 2 μ g p53-enhancer-luciferase plasmid (mdm2NA-Luc) and: (1) 8 μ g control vector pSRMSVtkNeo; (2) 5 and 8 μ g c-abl vector; (3) 5 and 8 μ g c-abl(K-R) vector; and (4) 5 and 8 μ g c-ablΔProl vector. Cells were also transfected with 2 μ g SV40-promoter-luciferase plasmid (pGL2-control vector; Promega) and 8 μ g c-abl vector. Luciferase activity was measured and normalized for protein concentration to that for control vector. **b**, MCF-7 cells were transfected with 8 μ g control, c-abl, c-abl(K-R) or c-ablΔProl vector. Cell lysates were immunoblotted with anti-p21 (WAF1, Ab-1; Oncogene Science) (top) and anti-PCNA (PC10; Santa Cruz Biotechnology) (bottom). **c**, Cell lysates from the transient transfectants were immunoprecipitated (IP) with anti-Cdk2 (M2, Santa Cruz Biotechnology), followed by histone H1 kinase assay (top) or immunoblotting (IB) with anti-Cdk2 (bottom).

METHODS. Transient transfections were done with Lipofectamine (GIBCO-BRL). Cells were collected 48 h after transfection. Luciferase was assayed with an enhanced luciferase assay kit (1800K, Analytical Luminescence). Histone H1 kinase was assayed as described¹⁶.

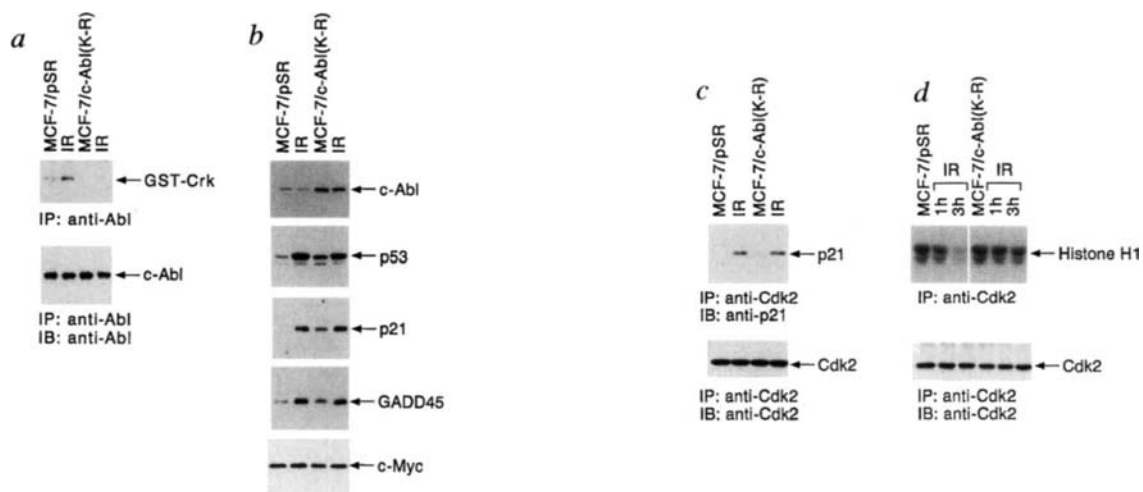


FIG. 2 c-Abl kinase activity regulates irradiation (IR)-induced inhibition of Cdk2. **a**, MCF-7 cells stably transfected with null pSR vector or *c-abl*(K-R) were treated with 5 Gy IR and collected after 3 h. Nuclei were isolated and nuclear proteins immunoprecipitated with anti-Abl (Ab-3; Oncogene Science) as described¹. Immune-complex kinase was assayed using a glutathione-S-transferase (GST)-Crk (120–225) fusion protein as substrate (top). Anti-Abl immunoprecipitates were also analysed by immunoblotting with anti-Abl (bottom). **b**, Lysates were immunoblotted with anti-Abl, anti-p53 (Ab-6; Oncogene Science), anti-p21, anti-GADD45 (AT-26; Santa Cruz Biotech-

nology) and anti-c-Myc (9E10; Santa Cruz Biotechnology). **c**, Lysates were immunoprecipitated with anti-Cdk2. Precipitates were analysed by immunoblotting with anti-p21 and anti-Cdk2. **d**, Anti-Cdk2 immunoprecipitates were analysed by histone H1 kinase assay (top) or immunoblotting with anti-Cdk2 (bottom).

METHODS. Human MCF-7 breast cancer cells (*p53*^{-/-})¹⁷ were treated with 5 Gy IR at room temperature with a Gammacell 1000 (Atomic Energy of Canada) with a ¹³⁷Cs source emitting at a fixed dose rate of 0.76 Gy min⁻¹. Lysates were prepared in 0.5% NP-40 lysis buffer as described¹.

(refs 7–10), by both wild-type c-Abl and c-Abl(K-R), Cdk2 was downregulated only in cells transfected with wild-type c-Abl (Fig. 1c), which correlates with the ability of wild-type c-Abl, but not c-Abl(K-R) or c-AblΔProl, to inhibit growth in fibroblasts^{2,3} and MCF-7 cells (data not shown). These findings suggest that the kinase activity and p53-binding domain of c-Abl are involved in downregulation of Cdk2 and thus in growth arrest, and that such effects are not exclusively mediated by activation of p21.

To determine the involvement of c-Abl in DNA-damage-dependent growth arrest, we prepared MCF-7 cells stably expressing the dominant-negative c-Abl(K-R)², which effectively inhibits the increase in c-Abl kinase activity induced by ionizing radiation on control cells (Fig. 2a). Irradiation of MCF-7/pSR or MCF-7/c-Abl(K-R) cells caused an increase in p53, p21 and GADD45, but not in c-Myc (which is not dependent on p53) (Fig. 2b); there was also increased binding of p21 to Cdk2 in MCF-7/pSR and in

MCF-7/c-Abl(K-R) cells (Fig. 2c). Although Cdk2 was downregulated in irradiated MCF-7/pSR cells, there was little effect on Cdk2 activity in MCF-7/c-Abl(K-R) cells (Fig. 2d). These results indicate that c-Abl kinase is not required for transactivation of p21 and GADD45 by p53 in response to DNA damage, but is required for Cdk2 downregulation.

The role of c-Abl in growth arrest induced by ionizing radiation was tested by the ability of irradiated MCF-7/pSR and MCF-7/c-Abl(K-R) cells to reduce the number of cells in S phase. Using bromo-deoxyuridine (BrdU) labelling and bivariate fluorescence-activated cell-sorting (FACS) analysis, we found that more irradiated MCF-7/pSR cells arrested in G1 compared with controls (Fig. 3a). Only 15% of cells were in S phase 24 h after receiving 5 Gy radiation compared with untreated cells. However, irradiated MCF-7/c-Abl(K-R) cells were less affected, with the S-phase population being 45% of untreated samples (Fig. 3a).

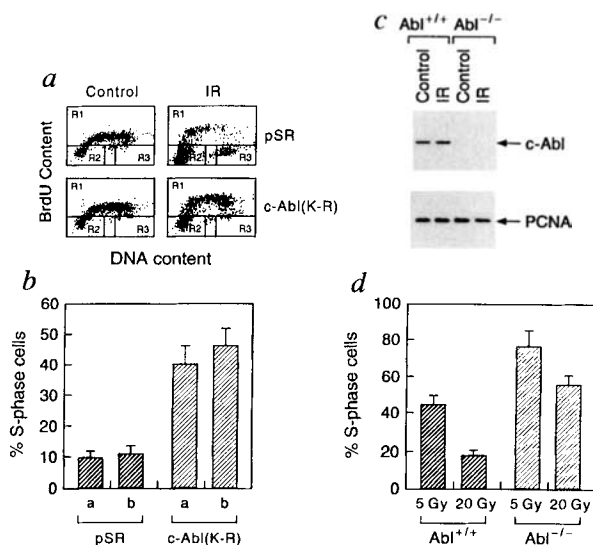


FIG. 3 c-Abl kinase activity regulates irradiation-induced growth arrest. **a**, Representative two-dimensional FACS analysis of MCF-7/pSR and MCF-7/c-Abl(K-R) cells after exposure to 0 or 5 Gy irradiation. Synchronized cells were irradiated and the percentage of cells in S phase was assessed at 24 h. Boxes labelled R1, R2 and R3 represent S, G1 and G2/M phase cells, respectively. **b**, Percentage of cells entering S phase after irradiation relative to control unirradiated cells. Results are expressed as the mean $\pm$ s.e.m. of 6 experiments for each of two (a, b) independently selected clones. **c**, C57BL/6 wild-type (*abl*^{+/+}) and *abl*^{-/-} MEFs were exposed 0 or 5 Gy and collected at 3 h. Cell lysates were immunoblotted with anti-Abl and anti-PCNA antibodies. **d**, *abl*^{+/+} and *abl*^{-/-} MEFs were exposed to 0, 5 or 20 Gy irradiation. Results (mean $\pm$ s.e.m. of 6 experiments) are expressed as the percentage of cells entering S phase after irradiation relative to unirradiated cells.

METHODS. MCF-7/pSR and MCF-7/c-Abl(K-R) cells were blocked at G1/S phase by a 12-h treatment with 10 μ M aphidicolin. Cells were then washed, irradiated and fed with fresh medium containing 10% FCS. Asynchronous populations of *abl*^{+/+} and *abl*^{-/-} MEFs¹¹ were exposed to IR and the percentage of cells in S phase assessed at 20 h. BrdU was added 30 min before collection. Cells were stained for DNA content with propidium iodide and for DNA synthesis with a fluorescein-conjugated anti-BrdU antibody (Boehringer Mannheim).

Similar results were obtained with two independently isolated MCF-7/c-Abl(K-R) clones (Fig. 3b). To confirm the link between c-Abl and radiation induced growth arrest, we studied mouse embryo fibroblasts (MEFs) deficient in c-Abl (*abl*^{-/-} mice with targeted disruption of the *c-abl* gene)¹¹ (Fig. 3c). Wild-type MEFs after 5 Gy radiation had an S-phase population that was 45% of that for untreated cells (Fig. 3d). By contrast, irradiated *abl*^{-/-} MEFs had more than 70% of cells in S phase compared with controls (Fig. 3d). Exposure of the wild-type and *abl*^{-/-} MEFs to 20 Gy radiation also demonstrated partial inhibition of arrest in G1 in the *abl*^{-/-} cells. These findings together indicate that the c-Abl kinase function is necessary for radiation-induced G1 arrest.

As c-Abl and p53 bind together *in vitro*⁴, we tested whether they were associated in irradiated cells. No p53 was detected in anti-Abl immunoprecipitates from MCF-7/pSR cells that had not been irradiated (Fig. 4a), but binding of c-Abl and p53 was evident at 30 min and maximal at 3–5 h after irradiation (Fig. 4a, and data not shown). The small amount of constitutive binding of c-Abl to p53 found in MCF-7/c-Abl(K-R) cells was stimulated by irradiation (Fig. 4a).

To investigate further the role of p53 in growth regulation by c-Abl, we transfected *p53*^{-/-} fibroblasts with wild-type c-Abl and c-Abl(K-R). Analysis of p21 levels and Cdk2 activity failed to reveal any effect with either vector (Fig. 4b, and data not shown). However, Cdk2 was downregulated in *p53*^{+/+} fibroblasts after transfection with wild-type c-Abl, but not c-Abl(K-R) (Fig. 4b). These results show that c-Abl kinase inhibits Cdk2 by a p53-dependent mechanism. We also investigated the effect on Cdk2 activity of irradiating *abl*^{-/-} cells and found that this did not decrease as much as in wild-type MEFs (Fig. 4c). In *p21*^{-/-} fibroblasts, it was found that p21 only partially mediates the action of p53 in arresting radiation-damaged cells in G1 (refs 12, 13). In support of our previous findings, expression of c-Abl(K-R) had no effect on Cdk2 activity in *p21*^{-/-} or *p21*^{+/+} cells (Fig. 4d); whereas transfection of wild-type c-Abl into *p21*^{-/-} cells inhibited Cdk2 (Fig. 4d) and growth (data not shown), indicating that these effects of c-Abl depend on p53 but not p21. It is possible, however, that growth arrest following transfection with c-Abl inhibits Cdk2 activity.

Our results obtained using a dominant-negative c-Abl mutant suggest that c-Abl regulates growth in response to genotoxic stress. Because c-Abl associates with the retinoblastoma protein Rb¹⁴ and overexpression of kinase-defective c-Abl abrogates Rb-induced growth arrest¹⁵, mutant c-Abl may prevent G1 arrest in response to DNA damage by inactivating Rb. But as the partial loss of G1 arrest in irradiated cells expressing c-Abl(K-R) is associated with a block in Cdk2 downregulation, the effect probably occurs upstream of Rb. We have demonstrated that c-Abl functions in the cellular response to DNA damage through p53-dependent pathways, confirmed by the failure of the c-AblΔProl mutant to bind p53⁴ and downregulate Cdk2 and arrest growth. We have also

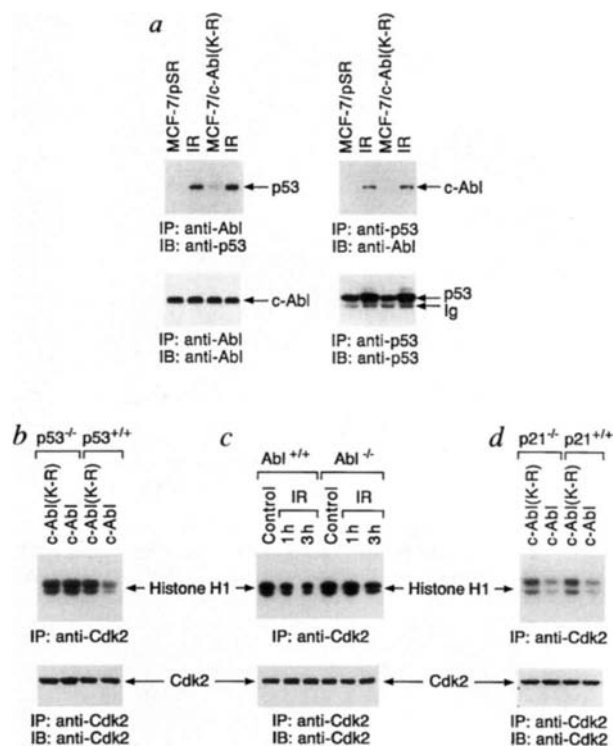


FIG. 4 c-Abl kinase downregulates Cdk2 activity by a p53-dependent, p21-independent mechanism. **a**, Lysates were immunoprecipitated with anti-Abl (left) or anti-p53 (right) and the precipitates analysed by immunoblotting with the indicated antibodies. Ig, immunoglobulin. **b**, MEFs (*p53*^{-/-}, *p53*^{+/+})¹⁸ were transfected with the *c-abl* or *c-abl*(K-R) vectors and collected at 48 h. **c**, MEFs (*abl*^{+/+}, *abl*^{-/-})¹¹ were exposed to 0 or 5 Gy and collected at the indicated times. **d**, MEFs (*p21*^{-/-}, *p21*^{+/+})¹² were transfected with the *c-abl* or *c-abl*(K-R) vectors and collected at 48 h. Cell lysates were immunoprecipitated with anti-Cdk2 and the precipitates assayed for histone H1 kinase activity (top) or immunoblotted with anti-Cdk2 (bottom) as described for Fig. 1.

shown that the kinase-defective c-Abl(K-R) mutant induces p53 transactivation but not G1 arrest, indicating that stimulation of p53 transactivation function by c-Abl is not sufficient for this response. Although c-Abl stimulates the transcriptional activity of p53, as measured by expression of its target gene *p21*, a different mechanism is used by p21 in growth arrest¹⁰ because c-Abl-mediated transactivation is kinase-independent and Cdk2 downregulation and growth arrest by the c-Abl kinase are p21-independent. These results support a novel c-Abl/p53 dependent, but p21-independent, mechanism for the regulation of Cdk2. □

Received 12 December 1995; accepted 12 June 1996.

1. Kharbanda, S. et al. *Nature* **376**, 785–788 (1995).
2. Sawyers, C. L., McLaughlin, J., Goga, A., Havlik, M. & Witte, O. *Cell* **77**, 121–131 (1994).
3. Mattioli, T., Jackson, P. K., Bchini-Hoof van Huijsduijn, O. & Picard, D. *Oncogene* **10**, 1325–1333 (1995).
4. Goga, A. et al. *Oncogene* **11**, 791–799 (1995).
5. Barak, Y., Gottlieb, E., Juven-Gershon, T. & Oren, M. *Genes Dev.* **8**, 1739–1749 (1994).
6. El-Deiry, W. et al. *Cell* **75**, 817–825 (1993).
7. Gu, Y., Turck, C. & Morgan, D. *Nature* **366**, 707–710 (1993).
8. Harper, J. W. et al. *Molec. Biol. Cell* **6**, 387–400 (1995).
9. Xiong, Y. et al. *Nature* **366**, 701–704 (1993).
10. Dulic, V. et al. *Cell* **76**, 1013–1023 (1994).
11. Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).

12. Deng, C., Zhang, P., Harper, J. W., Elledge, S. J. & Leder, P. *Cell* **82**, 675–684 (1995).
13. Brugarolas, J. et al. *Nature* **377**, 552–557 (1995).
14. Welch, P. J. & Wang, J. Y. *J. Cell* **75**, 779–790 (1993).
15. Welch, P. J. & Wang, J. Y. *J. Molec. cell. Biol.* **15**, 5542–5551 (1995).
16. Yuan, Z.-M., Kharbanda, S. & Kufe, D. *Biochemistry* **34**, 1058–1063 (1995).
17. Bartek, J., Vojtesek, B., Grand, R. J. A., Gallimore, P. H. & Lane, D. P. *Oncogene* **7**, 101–108 (1992).
18. Lowe, S. W., Jacks, T., Housman, D. E. & Ruley, H. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2026–2030 (1994).

ACKNOWLEDGEMENTS. We thank B. Mayer for early-passage *abl*^{-/-} MEFs, and P. Leder for *p21*^{-/-} fibroblasts. This investigation was supported by a PHS grant from the National Cancer Institute, DHHS.

CORRESPONDENCE and requests for materials should be addressed to (e-mail: donald_kufe@macmillan.fdcfl.harvard.edu).

An agent cleaving glucose-derived protein crosslinks *in vitro* and *in vivo*

Sara Vasan*, Xin Zhang*, Xini Zhang, Aphrodite Kapurniotu, Jürgen Bernhagen, Saul Teichberg†, John Basgen‡, Dilip Wagle*, David Shih*, Ihor Terlecky*, Richard Bucala§, Anthony Cerami, John Egan* & Peter Ulrich

The Picower Institute for Medical Research, 350 Community Drive, Manhasset, New York 11030, USA

* Alteon Inc. Ramsey, New Jersey 07446, USA

† North Shore University Hospital, Manhasset, New York 11030, USA

‡ University of Minnesota School of Medicine, Minneapolis, Minnesota 55455, USA

GLUCOSE and other reducing sugars react with proteins by a non-enzymatic, post-translational modification process called non-enzymatic glycosylation or glycation. The sugar-derived carbonyl group adds to a free amine, forming a reversible adduct which over time rearranges to produce a class of products termed advanced-glycation end-products (AGEs). These remain irreversibly bound to macromolecules and can covalently crosslink proximate amino groups^{1,2}. The formation of AGEs on long-lived connective tissue and matrix components accounts largely for the increase in collagen crosslinking that accompanies normal ageing and which occurs at an accelerated rate in diabetes^{3,4}. AGEs can activate cellular receptors and initiate a variety of pathophysiological responses⁵⁻⁹. They modify an appreciable fraction of circulating low-density lipoproteins preventing uptake of these particles by their high-affinity tissue receptors^{10,11}. Advanced glycation has also been implicated in the pathology of Alzheimer's disease^{12,13}. Because AGEs may form by a pathway involving reactive α -dicarbonyl intermediates^{1,2,14}, we investigated a potential pharmacological strategy for selectively cleaving the resultant glucose-derived protein crosslinks. We now describe a prototypic AGE crosslink 'breaker', *N*-phenacylthiazolium bromide (PTB), which reacts with and cleaves covalent, AGE-derived protein crosslinks. The ability of PTB to break AGE crosslinks *in vivo* points to the importance of an α -dicarbonyl intermediate in the advanced glycation pathway and offers a potential therapeutic approach for the removal of established AGE crosslinks.

The dominant structure of the crosslinking AGEs that form *in vivo* has been difficult to determine because of the structural heterogeneity of the principal crosslinks, the low abundance of individual crosslinking species, and the lability of AGEs to the conditions required for their purification^{1,2}. Model systems have identified one likely route to the formation of covalent, glucose-derived protein crosslinks¹⁴. This pathway proceeds by dehydration of the Amadori product (AP) in successive β -eliminations (Fig. 1a). Thus, loss of the 4-hydroxyl of the AP (molecule I in Fig. 1a) gives a 1,4-dideoxy-1-alkylamino-2,3-hexodiolose (AP-dione) (molecule II). Subsequent elimination of the 5-hydroxyl group gives a 1,4,5-trideoxy-1-alkylamino-

2,3-hexulos-4-ene (AP-ene-dione) (molecule III), which has been isolated as a triacetyl derivative of its 1,2-enol form¹⁵. Amadori diones, particularly the AP-ene-dione, would be expected to be highly reactive in protein crosslinking reactions by serving as targets for the addition of amine(Lys, His)- or sulphhydryl(Cys)-based nucleophiles, producing stable crosslinks of the form IV in Fig. 1a.

The possibility that a major pathway of glucose-derived crosslink formation proceeds through an AP-ene-dione intermediate suggested to us an experimental approach that would both test the occurrence of this pathway *in vivo* as well as effect the specific cleavage of the resultant α -dicarbonyl-based protein crosslinks. The carbon-carbon bond of α -diketones can be selectively cleaved, for example in reactions with certain thiazolium salts¹⁶. We prepared a series of thiazolium compounds to test their activity as bidentate nucleophiles in breaking the carbon-carbon bond between the two carbonyls of an AGE crosslink. Figure 1b shows a proposed scheme for the reaction of a prototypic α -dione cleaving agent, *N*-phenacylthiazolium bromide (PTB), with an AP-ene-dione-derived crosslink. The nucleophilic centres at the thiazolium-2 position and at the α -position of the *N*-substituent should each react with a carbonyl group, which would produce a five-membered ring and thereby convert the carbonyl carbons to a tetrahedral geometry and facilitate spontaneous cleavage at physiological pH. PTB was synthesized and tested for its ability to effect carbon-carbon bond cleavage of 1-phenyl-1,2-propanedione, a model α -dicarbonyl compound. Incubating PTB with 1-phenyl-1,2-propanedione produced the predicted fission product, benzoic acid, in an amount that was equivalent to the added PTB ('100%' in Fig. 2a).

Once glucose-derived addition products form on proteins, additional reactions will occur over time to produce covalent protein-protein crosslinks^{1,17,18}. We tested the ability of PTB to

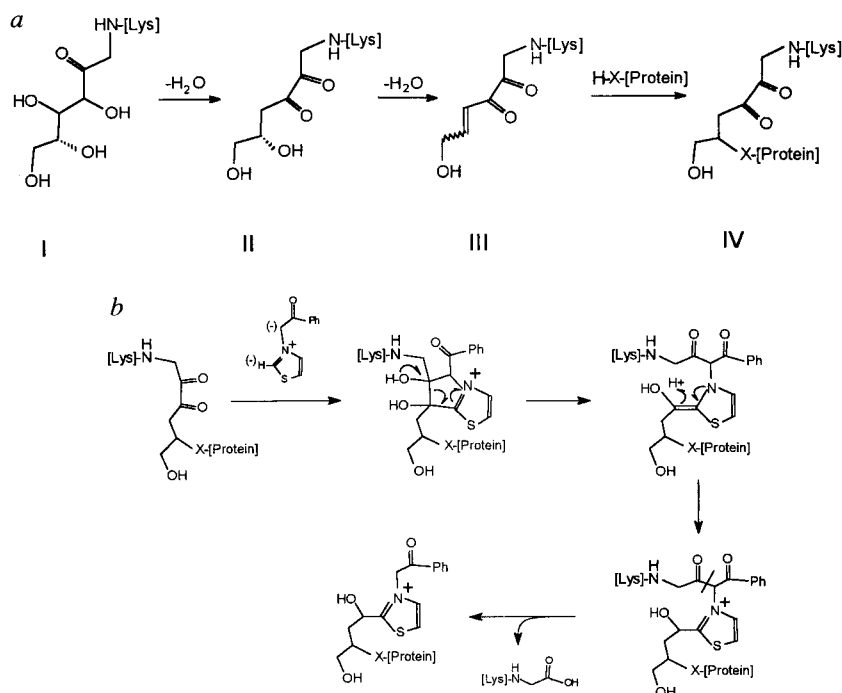
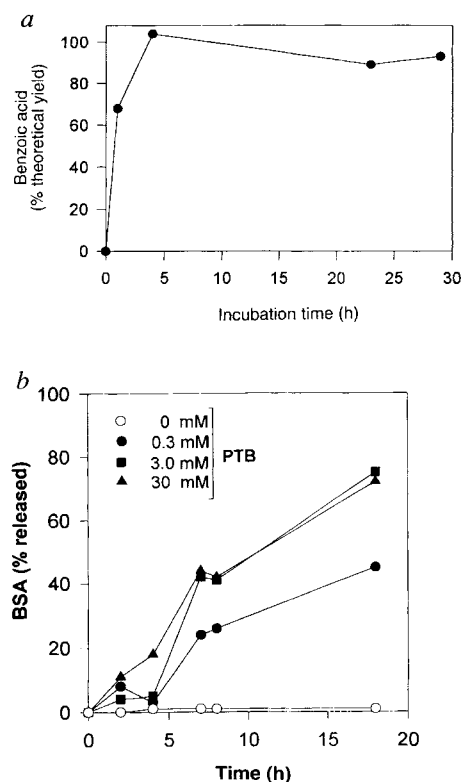


FIG. 1 Scheme for the formation of glucose-derived protein crosslinks from Amadori products and cleavage by a thiazolium-based AGE breaker. a, Successive dehydration by β -elimination of protein-bound (lysine) Amadori products (I) to AP-dione (II), AP-ene-dione (III), and reaction with a protein nucleophile (X-[Protein]) to form a stable protein-protein crosslink (IV). Molecule I exists predominantly in a pyranose form, and II, IV and the *cis* form of III may also prefer pyranose- or furanose-like cyclic hemiacetal structures¹⁷. b, Proposed reaction scheme for the cleavage of an AP-ene-dione-derived, protein-protein crosslink by *N*-phenacylthiazolium bromide (PTB).

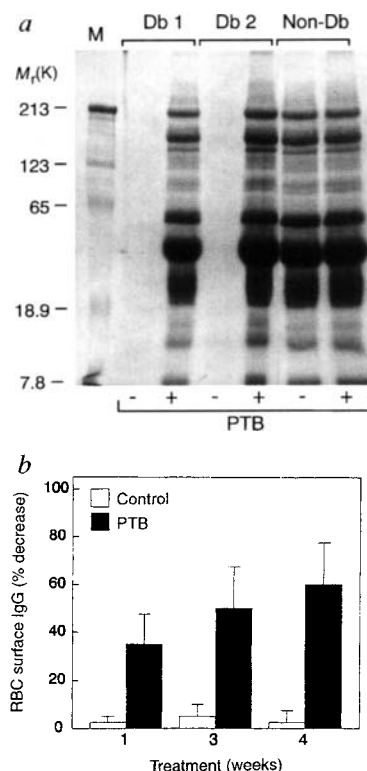
FIG. 2 PTB cleaves α -diketones and AGE-crosslinks *in vitro*. **a**, Benzoic acid formation from reaction of PTB with 1-phenyl-1,2-propanedione. PTB was prepared in 74% yield by heating a solution 1M in both phenacyl bromide and thiazole in ethanol at reflux for 2 h, cooling, filtering out the precipitate, and recrystallizing from aqueous 90% ethanol. Melting point, 222–223 °C (literature m.pt, 223–223.5 (dec.))²³. 1-Phenyl-1,2-propanedione (22 mM) was incubated with 10 mM PTB in 0.5 M sodium phosphate buffer (pH 7.4) containing 50% methanol under N₂ at 37 °C. Aliquots of the reaction mixture were removed at intervals and analysed by HPLC using a C18 reverse-phase column eluted with a gradient of 25–100% methanol. Benzoic acid was identified by the ¹H NMR spectrum and low- and high-resolution electron-impact mass spectra of the peak material isolated by preparative HPLC. The yield of benzoic acid was calculated by area integration (A_{254}) and interpolation from a standard curve obtained by injection of standard amounts of pure benzoic acid. **b**, PTB cleaves AGE-BSA–collagen crosslinks. AGE-BSA was prepared by incubating glucose with BSA for 12 weeks at 37 °C (ref. 24). After extensive dialysis to remove unreacted glucose, 1.0 μ g AGE-BSA was added to each well of a collagen-coated, 96-well microtitre plate (Biocoat, Collaborative Research). AGE-BSA was allowed to react with collagen for 4 h at 37 °C, after which wells were washed 3 times with PBS/0.05% Tween to remove unattached AGE-BSA. Test concentrations of PTB (50 μ l) dissolved in PBS were added to wells in triplicate and incubation continued at 37 °C for the indicated times. Wells then were washed extensively with PBS/0.05% Tween and the amount of BSA remaining attached to collagen quantified by direct ELISA using a rabbit polyclonal anti-BSA antibody, a horseradish peroxidase-linked goat anti-rabbit antibody, and H₂O₂ substrate containing ABTS (2,2'-azino-di-3-ethylbenzthiazoline sulphonic acid) as chromogen. The per cent breaking activity was calculated as $100 \times ((A_{410}, \text{PBS control}) - (A_{410}, \text{PTB})) / (A_{410}, \text{PBS control})$. Controls confirmed that PTB did not cause nonspecific dissociation of collagen from the microtitre plate surface. Addition of PTB (0.003–3 mM) to wells for 7 h at 37 °C followed by washing and incubation with AGE-BSA did not affect the ability of AGE-BSA to then attach to the collagen-coated surface, as detected by direct BSA ELISA. PTB did not react either with AGE-BSA to modify or obscure the BSA epitopes recognized by anti-BSA antibody. AGE-BSA (40 μ g ml⁻¹) was incubated with PTB (0.003–3 mM) in PBS for 7 h at



37 °C, dialysed, and subjected to BSA-specific ELISA. No significant differences in immunoreactivity were observed when PTB-treated and untreated preparations were compared (data not shown).

FIG. 3 PTB cleaves AGE crosslinks that form *in vivo*. **a**, PTB treatment *in vitro* decreases AGE crosslinking in diabetic, rat tail-tendon collagen. Data are representative of collagen from two diabetic animals (Db1, Db2) and one non-diabetic control (non-Db). Lane M, relative molecular mass markers. **b**, PTB treatment decreases IgG crosslinked to the red-blood-cell (RBC) surface.

METHODS. **a**, Diabetes was induced in male Lewis rats (150–175 g) by i.p. injection of streptozotocin (65 mg kg⁻¹) and confirmed after one week by plasma glucose measurement (≥ 250 mg dl⁻¹). Thirty-two weeks later, the rats were killed and collagen isolated from their tail-tendon fibres using a standard protocol²⁵. The insoluble collagen then was treated with cyanogen bromide²⁶ and the hydroxyproline content measured²⁷. Aliquots containing 1 μ g equivalent of hydroxyproline were run on SDS–PAGE under reducing conditions and stained with Coomassie blue. **b**, IgG crosslinked to the RBC surface was determined in an anti-IgG ELISA adapted for use with cellulose-ester-membrane-sealed 96-well microtitre plates (Multiscreen-HA, Millipore). Heparinized blood was washed three times with PBS; the packed RBC were diluted 1:250–1:500 in PBS. Membrane-containing wells were first blocked with 0.3 ml Superblock (Pierce), then washed with 0.3 ml PBS/0.05% Tween, followed by 0.1 ml PBS. RBCs were gently vortexed and 50- μ l aliquots pipetted into wells. Cells were then washed 3 times with PBS, once with Tris-buffered saline, and 0.1 ml *p*-nitrophenyl phosphate substrate was added (1 mg ml⁻¹ in 0.1 M diethanolamine buffer, pH 9.5). By this technique, the A_{410} of non-diabetic red cells was 0.10 ± 0.02 and the A_{410} of diabetic red cells was 0.57 ± 0.06 ($n = 4$; $P < 0.0002$). The ability of PTB to reduce RBC-surface IgG *in vitro* was evaluated as follows. RBCs from diabetic rats were washed and 0.1-ml aliquots incubated overnight at 37 °C with 1 ml PTB in PBS. Control incubations contained RBCs and PBS alone. At the end of the reaction, RBCs were assayed for RBC-IgG; per cent decrease was calculated as $100 \times ((A_{410}, \text{PBS control}) - (A_{410}, \text{PTB})) / (A_{410}, \text{PBS control})$. Diabetic rats were treated for up to 4 weeks with PTB (10 mg kg⁻¹ q.d. by oral gavage) or saline as control ($n = 4$ –6 rats per group). At intervals, blood was collected from tail veins into heparinized tubes, washed 3 times with 10 vol PBS, and assayed for surface IgG. Values are means $\pm$ s.d. of the per cent decrease with respect to day 0. *P* values were calculated by the Student's *t*-test, independent variable.



disrupt the crosslinks that form when AGE-modified bovine serum albumin (AGE-BSA) is allowed to react with unmodified native collagen (Fig. 2b legend). As shown in Fig. 2b, PTB produced a concentration-dependent release of BSA from pre-formed AGE-BSA–collagen complexes. These data support the concept that a significant portion of the crosslinks that form by reaction of AGE-BSA with collagen contain an α -diketone or a closely related, linking structure that is susceptible to cleavage by thiazolium-based nucleophiles such as PTB. Whether PTB remains stably linked to one partner of the original protein–protein crosslink remains to be determined.

We next tested the ability of PTB to cleave AGE crosslinks that form *in vivo*. Progressive advanced glycation of long-lived matrix proteins occurs under the excess glucose conditions present in diabetes and increases collagen crosslinking, resistance to proteolytic and chemical digestion, and the rigidity of connective tissue^{3,4,18}. We isolated collagen from the tail tendons of rats that had been diabetic for 32 weeks and treated it with PTB *in vitro* before cyanogen bromide digestion and gel electrophoretic analysis. As expected, the increase in collagen crosslinking produced by advanced glycation resulted in a marked decrease in the susceptibility of tail tendon collagen to cyanogen bromide digestion. Thus, few collagen peptides were visualized after SDS–polyacrylamide gel electrophoresis (SDS–PAGE; Fig. 3a). By contrast, diabetic collagen that had been treated with PTB before cyanogen bromide digestions produced a peptide pattern almost indistinguishable from untreated, non-diabetic collagen.

The products of advanced glycation are formed *in vivo* on a variety of blood constituents, including haemoglobin¹⁹, the lipid and apolipoprotein components of low-density lipoproteins (LDLs)^{10,11}, β_2 -microglobulin²⁰ and immunoglobulins²¹. We isolated red cells from diabetic rats and measured the amount of

surface-bound IgG after treatment *in vitro* with PTB. About five times as much IgG was bound to the cell surface of diabetic red cells compared with non-diabetic red cells when analysed by enzyme-linked immunosorbent assay (ELISA) using an alkaline phosphatase-linked, rabbit anti-rat IgG primary antibody. Treating diabetic red cells with 1 mM PTB overnight (37 °C) decreased the amount of detectable surface IgG by $83 \pm 15\%$.

We next tested the efficiency of PTB as an AGE-crosslink breaker *in vivo*. Rats with established diabetes were administered with daily intraperitoneal (i.p.) injections of PTB (10 mg kg^{-1}) and blood samples collected at intervals for erythrocyte cell-surface IgG determinations. When compared with the control group, animals treated with PTB had markedly less IgG bound to their erythrocytes (Fig. 3b). PTB treatment decreased cell-surface IgG content by $36 \pm 13\%$ after one week ($P < 0.0005$), $47 \pm 19\%$ after three weeks ($P < 0.0003$), and $60 \pm 18\%$ after four weeks ($P < 0.001$).

Immunoreactive AGEs are present in increased amounts in the brain tissue of patients with Alzheimer's disease, localize to the amyloid plaque deposits, and can act as efficient nucleation 'seeds' *in vitro* for the polymerization and progressive insolubilization of the β -amyloid peptide (β AP)^{12,13}. We therefore tested whether PTB could disaggregate the β -amyloid that forms by AGE-modification of native, monomeric β AP¹². Fibrillar aggregates of AGE-modified β AP were radioiodinated, incubated with PTB, and then subjected to SDS–PAGE under reducing conditions. As expected, β AP untreated with PTB migrated as high-molecular-weight aggregates at the very top of the polyacrylamide running gel (Fig. 4a). Treatment of β AP with PTB *in vitro*, however, produced a markedly different electrophoretic pattern: in addition to high-molecular-weight aggregates at the top of the gel, significant amounts of smaller β AP forms were present. The

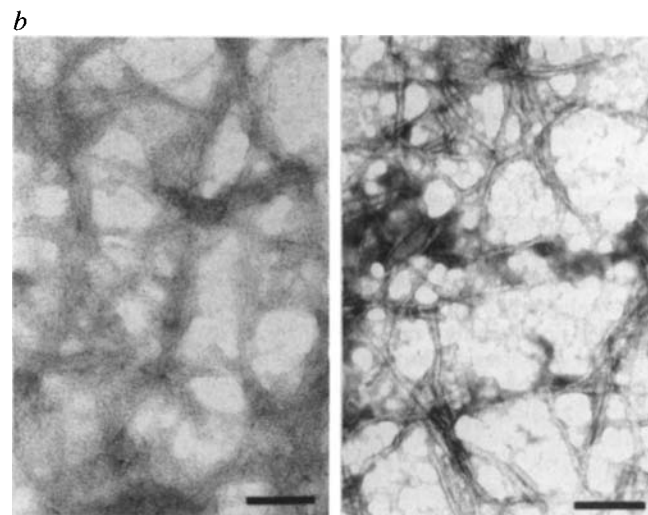
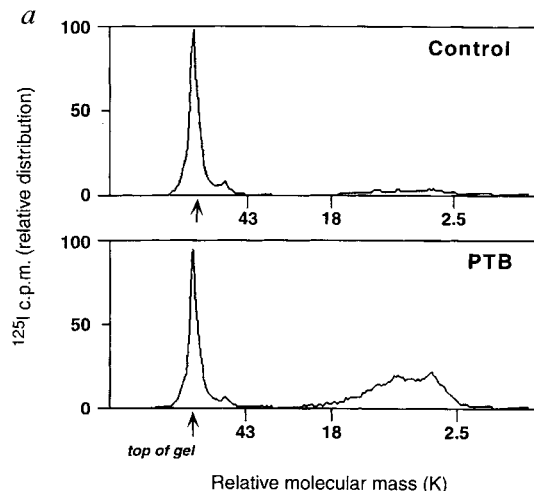


FIG. 4 PTB disaggregates β -amyloid fibrils *in vitro*. a, SDS–PAGE and radiometric scanning analysis of ^{125}I -labelled, β -amyloid peptide (β -AP) aggregated *in vitro* by AGE modification. Top, ^{125}I - β -AP aggregates incubated with PBS (control). Bottom, ^{125}I - β -AP aggregates incubated with 10 mM PTB (PTB). b, Electron microscopic analysis of AGE- β -AP treated with PTB (left) or saline as control (right).

METHODS. a, β -AP (amino acids 1–40) from Bachem (Torrance, CA) was glycated *in vitro* and dialysed to produce high-molecular-weight AGE-modified β -AP 'seeds' as described¹². AGE- β -AP was radioiodinated using the IODO-beads iodination reagent (Pierce). Briefly, 50 μl AGE- β -AP (25 μM) was added to a solution of 215 μl PBS (pH 6.9) containing a washed IODO-bead and 15 μl ^{125}I (100 mCi ml^{-1} , NEN Dupont). Radioiodination was stopped after 15 min by removing the bead, and free iodine removed by dialysis. For SDS–PAGE, 40 μl ^{125}I -labelled AGE- β -AP aggregate in 10 mM sodium phosphate buffer (pH 7.4) was incubated either alone or with 20 mM PTB for 24 h at 37 °C. Samples then were mixed with

an equal volume of 2 $\times$ concentrated Laemmli sample buffer, heated to 100 °C for 10 min, and electrophoresed in a 16% gel under reducing conditions. Protein molecular mass markers ranged from 2.5K–43K (Gibco-BRL). The gel was fixed in 20% trichloroacetic acid for 16 h and analysed by ^{125}I scanning with a Packard Instant Imager. The relative distribution of ^{125}I -labelled β -AP was measured in each lane, which was gated and scanned individually. b, 20 μl AGE- β -AP aggregate in 10 mM sodium phosphate (pH 7.4) was incubated either alone or with 20 mM PTB for 24 h at 37 °C. A 10- μl drop of each sample was placed on a formvar-coated, carbon-stabilized 200-mesh copper grid for 3 min and the excess liquid removed slowly with filter paper. A drop of 2% (w/v) uranyl acetate was then placed on the grids for 1 min, the excess removed, and the sample allowed to dry. Preparations were examined in a JEOL 100 CXII transmission electron microscope at 80 kV (magnification of each panel: 142,500 $\times$; scale bar, 100 nm).

relative extent of amyloid disaggregation appeared to be less than the decrease in crosslinking induced by PTB in tail-tendon collagen or red-cell-surface IgG. This may reflect a lesser contribution by AGE crosslinks to the physicochemical changes underlying β AP amyloid formation¹². Electron microscopy revealed marked differences in the fibrillar structure of β AP aggregates before and after treatment with PTB (Fig. 4b). In control (untreated) preparations, β -amyloid was organized into dense fibrillar aggregates. The tubular-like fibrils were occasionally paired and showed a relatively uniform diameter of 9 ± 3 nm, a size consistent with that of β -amyloid fibrils *in vivo*²². By contrast, the preparation of β AP treated with PTB showed fibrils that were

less dense, less uniformly organized, and more filamentous in form.

Our results support the hypothesis that a significant portion of the AGE crosslinks that form from glucose involve an α -diketone moiety that is susceptible to cleavage by thiazolium compounds such as PTB. Direct evidence for the formation of α -diketone crosslinks *in vivo* must await the isolation and complete characterization of these products. Although decreasing the AGE content of tissues may not necessarily ameliorate tissue pathology, this work does provide a rational basis for the design of AGE-cleaving agents that may ultimately be of broad clinical application. □

Received 1 May; accepted 24 May 1996.

- Njoroge, F. G. & Monnier, V. M. *Progr. clin. Biol. Res.* **304**, 85–107 (1989).
- Bucala, R. & Cerami, A. *Adv. Pharmac.* **23**, 1–34 (1992).
- Monnier, V. M., Kohn, R. R. & Cerami, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 583–587 (1984).
- Monnier, V. M. *et al.* *New Engl. J. Med.* **314**, 403–408 (1986).
- Vlassara, H., Brownlee, M., Manogue, K. R., Dinarello, C. & Pasagian, A. *Science* **240**, 1546–1548 (1988).
- Esposito, C., Gerlach, H., Brett, J., Stern, D. & Vlassara, H. *J. exp. Med.* **174**, 1387–1407 (1989).
- Vlassara, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **89**, 12043–12047 (1992).
- Doi, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **89**, 2873–2877 (1992).
- Vlassara, V., Fuh, H., Donnelly, T. & Cybulsky, M. *Molec. Med.* **1**, 447–456 (1995).
- Bucala, R., Makita, Z., Koschinsky, T., Cerami, A. & Vlassara, H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6434–6438 (1993).
- Bucala, R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 9441–9445 (1994).
- Vitek, M. P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 4766–4770 (1994).
- Smith, M. A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 5710–5714 (1994).
- Chen, H.-J. & Cerami, A. *J. Carbohydr. Chem.* **12**, 731–742 (1993).

- Estendorfer, S., Ledl, F. & Severin, T. *Angew. Chem. int'l Ed. Engl.* **29**, 536–537 (1990).
- Vovk, A. I., Murav'eva, I. V. & Yasnikov, A. A. *Ukr. Khim. Zh.* **51**, 521 (1985).
- Ledl, F. & Schleicher, E. *Angew. Chem.* **29**, 565–593 (1990).
- Eble, A. S., Thorpe, S. R. & Baynes, J. W. *J. biol. Chem.* **258**, 9406–9412 (1983).
- Makita, Z. *et al.* *Science* **258**, 651–653 (1992).
- Miyata, T. *et al.* *J. clin. Invest.* **92**, 1243–1252 (1993).
- Danze, P. M., Tarjoman, A., Rousseaux, J., Fossati, P. & Dautrevaux, M. *Clin. chim. Acta* **166**, 143–153 (1987).
- Cotra, R. S., Kumar, V. & Robbins, S. L. *Robbins Pathologic Basis of Disease* 5th edn, 231–238 (Saunders, Philadelphia, 1994).
- Ukai, O. K. *et al.* *Yakugaku Zasshi* **63**, 296 (1943).
- Makita, Z., Vlassara, H., Cerami, A. & Bucala, R. *J. biol. Chem.* **267**, 5133–5138 (1992).
- Bochantin, J. & Mays, L. L. *Exp. Gerontol.* **16**, 101–106 (1981).
- Itakura, M. *et al.* *Life Sci.* **49**, 889–897 (1991).
- Stegeman, H. & Stadler, K. *Clin. chim. Acta* **18**, 267–273 (1967).

ACKNOWLEDGEMENTS. We are grateful to M. Yamin for his careful reading of the manuscript.

CORRESPONDENCE and requests for materials should be addressed to R.B. (e-mail: rbucala@picower.edu).

Structure of *Taq* polymerase with DNA at the polymerase active site

Soo Hyun Eom*, Jimin Wang* & Thomas A. Steitz*†‡

Departments of * Molecular Biophysics & Biochemistry and † Chemistry, ‡ Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520-8114, USA

THE DNA polymerase from *Thermus aquaticus* (*Taq* polymerase) is homologous to *Escherichia coli* DNA polymerase I (Pol I) and likewise has domains responsible for DNA polymerase and 5' nuclease activities^{1,2}. The structures of the polymerase domains of *Taq* polymerase and of the Klenow fragment (KF) of Pol I are almost identical, whereas the structure of a vestigial editing 3'–5' exonuclease domain of *Taq* polymerase that lies between the other two domains is dramatically altered, resulting in the absence of this activity in the thermostable enzyme². The structures have been solved for editing complexes between KF and single-stranded DNA^{3,4} and for duplex DNA with a 3' overhanging single strand⁵, but not for a complex containing duplex DNA at the polymerase active-site. Here we present the co-crystal structure of *Taq* polymerase with a blunt-ended duplex DNA bound to the polymerase active-site cleft; the DNA neither bends nor goes through the large polymerase cleft, and the structural form of the bound DNA is between the B and A forms. A wide minor groove allows access to protein side chains that hydrogen-bond to the N3 of purines and the O2 of pyrimidines at the blunt-end terminus. Part of the DNA bound to the polymerase site shares a common binding site with DNA bound to the exonuclease site, but they are translated relative to each other by several ångströms along their helix axes.

Early models⁶ of primer–template DNA at the polymerase active site had the DNA running through the cleft that was a distinctive feature of the KF structure. Numerous considerations,

including the structure of an editing complex between KF and duplex DNA, as well as the location of completely conserved residues whose mutation affected polymerase activity⁷, led to the conclusion that the duplex primer–template DNA approached the polymerase active site from the 3'–5' exonuclease domain side of the cleft and perhaps was bent upon going through the cleft⁵. Based on comparisons between the structure of rat DNA polymerase- β (Pol- β) complexed with primer–template DNA and those of KF and HIV reverse transcriptase, it was suggested⁸ that the primer–template approaches KF from the side opposite the 3'–5' exonuclease domain, a conclusion that was disputed⁹. The crystal structure we describe here confirms the conclusion that the primer–template approaches the polymerase domain from the 3'–5' exonuclease side of the cleft.

Taq polymerase was co-crystallized with two 16-mers of DNA that form 8 base pairs of duplex and 8 nucleotides of non-complementary single strands (Fig. 1a). Crystals diffract to 3.0 Å resolution at -165°C using synchrotron radiation and are essentially isomorphous with those of the apo enzyme². An electron density map (Fig. 1b), calculated from the protein coordinates alone but improved by cycles of density modification¹⁰, shows good density for the bases and backbone of the duplex portion of the DNA.

The duplex DNA binds with its blunt end in the polymerase active site, packed against the O α -helix of the fingers domain and with its minor groove at the terminus interacting with side chains emanating from the palm domain (Fig. 2a). The DNA helix axis is roughly coincident with that of duplex DNA when bound in editing mode⁵ and does not pass through the polymerase cleft (Fig. 2b). The 3' hydroxyl of the primer strand is near to the three carboxylates (Asp 785[882], Glu 786[883], Asp 610[705]) that form the catalytic core of the polymerase active site, presumably by binding to divalent metal ions^{7,8,11}. (Amino-acid sequence numbers of KF are given in brackets). The side chains of Tyr 671[766] and Phe 667[762], which are known to be important from their conservation in the Pol I family of polymerases^{12,13} and from mutagenesis studies^{13–16}, are interacting with the bases at the blunt end of the duplex (Fig. 2c). Highly conserved amino-acid residues at the tip of the thumb (Asn 485[581], Ser 515[610], and

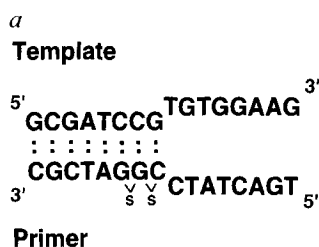
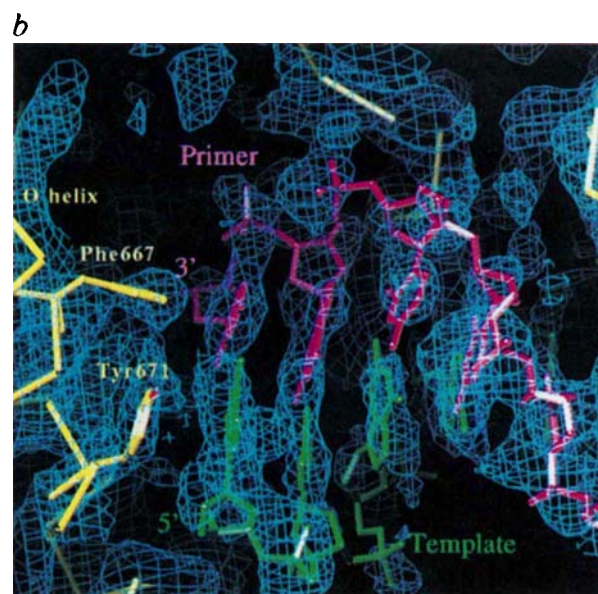


FIG. 1 Blunt-end DNA duplex co-crystallized with *Taq* polymerase. **a**, The DNA sequences used form 8 base pairs (bp) of duplex. Two nucleotides in the primer strand contained thiophosphates to prevent cleavage by the 5' nuclease domain. The complex was crystallized at 22 °C in hanging drops containing 100 mM sodium citrate, 4% (w/v) polyethylene glycol 8000, 1 mM ZnSO₄, and 0–400 mM KCl at pH 6.0. Co-crystals were flash-frozen after transferring to a stabilizing solution containing 50 mM sodium citrate, 10% (v/v) glycerol, 15% (v/v) ethylene glycol, 100 mM KCl, 0.4% (w/v) β -octyl glucoside, and 10% (w/v) polyethylene glycol 8000 for 2 min. The co-crystals are isomorphous with those of the apo enzyme², belonging to space group $P3_12_1$ and having cell dimensions of $a = b = 107.6$, $c = 170.1$ Å. A data set to 3.0 Å resolution was collected at the X12C beam line at BNL ($\lambda = 1.00$ Å) and processed with DENZO and SCALEPACK with an overall merging R -factor of 8.8% (ref. 26). The crystallographic R -factor was 34% after rigid-body refinement of the protein coordinates alone and dropped to 28% after tightly restrained positional refinement using X-PLOR (ref. 27). Phases calculated from the refined protein structure were combined with the experimental phases of the apo *Taq* polymerase using SIGMAA weighting²⁸ to 3.0 Å resolution. Electron density maps calculated with these combined phases showed good density for the DNA bases but poor density for the backbone. To improve the low-resolution phases, which typically are of poor quality when calculated from the model, a procedure using cycles of density modification¹⁰ to extend phasing from medium to low resolution was employed to improve the low-resolution phases. **b**, A 3-Å-resolution



solvent-flattened electron density map resulting from the density modification procedure to improve the phases calculated from the coordinates of the protein in atoms alone shows good density for the bases as well as the backbone. The refined structure is superimposed. The 3' base of the primer strand (pink) is interacting with Phe 667[762] on the O helix of the fingers domain and the 5' base of the template strand (green) is interacting with Tyr 671[766]. Additional refinement including 8 bp of DNA reduced the R -factor to 23% and the free R -factor to 30%, with an r.m.s. deviation of bond lengths from ideality of 0.009 Å and bond angles of 1.5°, respectively.

Lys 540[635]) interact directly or indirectly with the DNA backbone across the minor groove (Fig. 2c). The small domain at the tip of the thumb lies over the 'top' of the DNA that is enclosed on three sides by the protein.

Most of the amino-acid residues in the polymerase domain that are highly conserved among the Pol I family^{12,13} are in contact with or are very close to the bound duplex DNA (Fig. 2c), consistent with this complex being important in DNA synthesis. Seven absolutely conserved residues (Arg 573[668], Glu 655[710], Tyr 671[766], Asn 750[845], Gln 745[849], His 784[881] and Asp 785[882]) contact the DNA (Fig. 2c). However, while the three catalytically important carboxylates are near to the 3' nucleotide, they are too far away by 2–3 Å for metal ions bound to them to contact the 3' hydroxyl group of the primer strand. The 3' end of the duplex DNA observed here may lie further from the carboxylates than would be expected in a catalytically productive complex, due either to the lack of bound divalent metal ions or constraints on DNA positioning by crystal lattice interactions, or because it represents a distinct step in the polymerase reaction before or after the phosphoryl transfer step.

Comparison of the structure of the duplex DNA in this complex with those of A- and B-form DNAs shows that it has characteristics of both A and B form but is not identical to either (Fig. 3). Like A-form DNA, which has a groove width of 11.0 Å, the Pol I-DNA has a significantly wider minor groove (9.2 Å) than B-form DNA (5.7 Å) (ref. 17). Unlike A-form DNA, however, the base pairs are not highly tilted relative to the helix axis, nor are they significantly displaced outwards from the helix axis, which results in the bound DNA having a deeper minor groove than in A-form DNA. The DNA bound to Pol- β (ref. 8) has a similarly widened minor groove with characteristics of both A- and B-form DNA (Fig. 3).

Because the minor groove is wider than that of B-form DNA and faces the palm domain, it is accessible to protein side chains that can hydrogen-bond to the O2 atoms of pyrimidines and the N3 atoms of purines (Fig. 2c). These base hydrogen-bond accep-

tors lie in precisely the same place in the minor groove for all four base pairs¹⁸. It has been pointed out^{19–21} that as these functional groups are invariant with sequence and differently positioned for non-Watson-Crick base pairs, their specific recognition and positioning by protein side chains may be an important mechanism by which polymerases enhance the fidelity of Watson-Crick base pairing and correct nucleotide incorporation at the polymerase active site.

This *Taq* polymerase-DNA co-crystal structure confirms the view⁹ that the Pol I family of polymerases and HIV reverse transcriptase²² bind primer-template to the polymerase active site located in the 'palm' subdomain in similar ways. Superposition of the C α coordinates of the palm domains of these two polymerases puts the primer-template DNAs bound to the two enzymes in very similar orientations. The primer-template complex of *Taq* polymerase can be compared with that of DNA Pol- β by superimposing the bound DNA duplexes, as previously suggested⁹. This comparison shows many analogous interactions between the bound DNAs and these two non-homologous polymerases. In all three polymerases, the three catalytically important carboxylates of the palm domains are near the 3'-OH of the primer. Structurally unrelated helical protrusions (called 'thumb' in RT²³ and KF³, and fingers in Pol- β ⁸) interact across the minor groove of the DNA product. The blunt end of a primer-template abuts a similarly positioned α -helix in the fingers domain of *Taq* polymerase and a non-homologous domain in Pol- β .

Comparison of the position of duplex DNA bound to the polymerase active site with that of DNA bound to the exonuclease active site (Fig. 2b) shows that although they share a common duplex DNA-binding site, the duplexes are translated relative to each other by several ångströms along their helix axes, resulting in different DNA-backbone interactions with the protein (Fig. 2c). It would appear that the duplex DNA binding site is constructed such that the DNA can slide by translation and may not be constrained to move only in a screw motion defined by the DNA backbone, a most surprising conclusion.

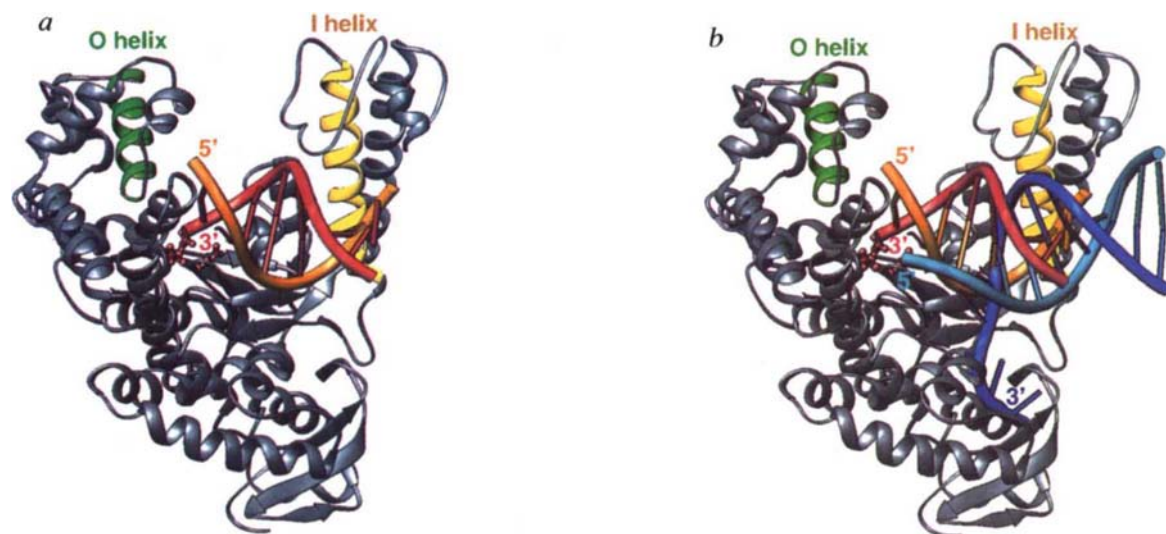
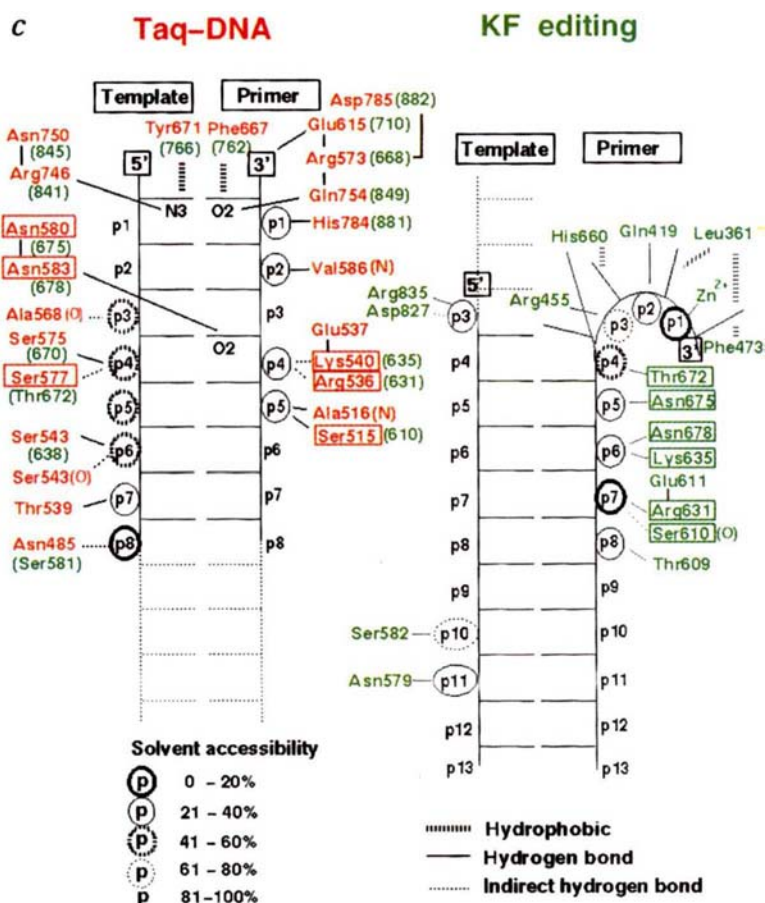
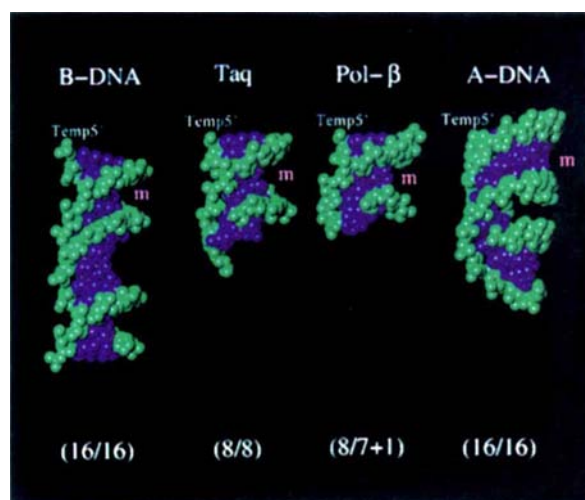


FIG. 2 The interaction of duplex DNA with *Taq* polymerase. *a*, Duplex portion of DNA on a ribbon representation of the KF portion of *Taq* polymerase. The O helix is in green, the I helix in yellow and the catalytically essential carboxylates, Asp 610[705], Asp 785[882] and Glu 786[883], are shown near the 3' terminus. The template strand is orange and the primer strand is red. *b*, Superposition of DNAs bound in editing and in polymerizing modes. To orient the two DNAs, the polymerase domains of the KF and *Taq* polymerase complexes were superimposed and the DNA from the editing complex added to the *Taq* polymerase-DNA complex. Particularly striking is the difference in the positions of the primer backbones relative to the loop at the tip of the I helix. The template strand of editing-complex DNA is in cyan and its primer strand is in blue. *c*, Diagrams of the interactions between DNA and protein in the *Taq* polymerase-DNA and KF editing complexes showing that many of the same side chains interact with different phosphates in the two complexes. The names of *Taq* polymerase side chains are red and those of KF are green. Backbone carbonyl oxygen and amide nitrogen atoms are in parentheses. Thick dotted lines are hydrophobic interactions; thin lines, hydrogen bonds; thin dotted lines, indirect hydrogen bonds that might be through water molecules. Solvent accessibilities of the phosphate groups have been calculated using the program ACCESS²⁹. Thick circles are 0–20% solvent-accessible; thin circles, 21–40%; thick dotted circles, 41–60%; thin dotted circles, 61–80%; blank, 81–100%. Side chains responsible for interactions and which interact in both complexes are boxed. The low solvent-accessibility of P5 of the template strand of the *Taq* polymerase-DNA complex results from Asp 578 (OD), Asp 578 (C α), Pro 579 (C α) and Asn 580 (N) lying within 4.5 Å. The coordinates used⁵ to evaluate the interactions with the editing DNA were derived from a 3.2-Å-resolution map. As not all of the backbone and side-chain atoms at the tip of the thumb were located in this complex, the protein protection of the



template strand is incomplete in this analysis. A 2.4-Å-resolution refinement of another editing complex is in progress (J. Jäger, D. Cahill and T.A.S., unpublished results).

FIG. 3 Comparison of the structures of the duplex DNAs bound to *Taq* and Pol- β with those of A- and B-form DNAs. The coordinates of DNA in the Pol- β complex was taken from the Protein Data Bank⁸. The helix axes have been tilted by 32° to show the major and minor (indicated by m) groove. Bases are shown in blue; backbones are in green. The lengths of the template–primer strands are given in parentheses.



Comparison of these two complexes raises the question of how the 3' terminus gets from one active site to another during the process of editing. It has been proposed^{3,24}, and supporting evidence obtained²⁵, that editing out misincorporated nucleotides is accomplished by competition for the 3' end of the primer strand between the polymerase active site, which binds base-paired duplex DNA, and the exonuclease active site, which binds single-stranded DNA. As mismatched base pairs destabilize duplex DNA and promote the formation of single-stranded DNA, a mismatched primer terminus has a greater propensity

to bind to the exonuclease site and be hydrolysed. The primer strand at its 3' terminus lies between the template strand and the protein, and it appears not to be physically possible to move it to the exonuclease active site without the DNA translocating back. Whereas the DNA would not have to dissociate from the enzyme to go from one active site to another (indeed, in the homologous T7 DNA polymerase–thioredoxin enzyme it does not³⁰), it appears that the 3' end would have to dissociate from the polymerase active site and slide before entering the exonuclease active site. □

Received 19 April; accepted 29 May 1996.

1. Lawyer, F. C. et al. *J. biol. Chem.* **264**, 6427–6437 (1989).
2. Kim, Y. et al. *Nature* **376**, 612–616 (1995).
3. Freemont, P. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8924–8928 (1988).
4. Beese, L. S. & Steitz, T. A. *EMBO J.* **10**, 25–33 (1991).
5. Beese, L. S., Derbyshire, V. & Steitz, T. A. *Science* **260**, 352–355 (1993).
6. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. *Nature* **313**, 762–766 (1985).
7. Joyce, C. M. & Steitz, T. A. *Rev. Biochem.* **63**, 777–822 (1994).
8. Pelletier, H. et al. *Science* **264**, 1891–1903 (1994).
9. Steitz, T. A., Smerdon, S. J., Jaeger, J. & Joyce, C. M. *Science* **266**, 2022–2025 (1994).
10. Cowtan, K. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34–38 (1994).
11. Steitz, T. A. *Curr. Opin. struct. Biol.* **3**, 31–38 (1993).
12. Braithwaite, D. K. & Ito, J. *Nucleic Acids Res.* **21**, 787–802 (1993).
13. Astatke, M., Grindley, N. D. F. & Joyce, C. M. *J. molec. Biol.* **270**, 1945–1954 (1995).
14. Polesky, A. H., Steitz, T. A., Grindley, N. D. F. & Joyce, C. M. *J. biol. Chem.* **265**, 14579–14591 (1990).
15. Polesky, A. H., Dahlberg, M. E., Benkovic, S. J., Grindley, N. D. F. & Joyce, C. M. *J. biol. Chem.* **267**, 8417–8428 (1992).
16. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **92**, 6339–6343 (1995).
17. Conner, B. N. et al. *Nature* **295**, 294–299 (1982).
18. Seeman, N. C., Rosenberg, J. M. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 804–808 (1976).
19. Steitz, T. A. et al. *Cold Spring Harbor Symp. quant. Biol.* **52**, 465–471 (1987).

20. Steitz, T. A. in *Biological Organization: Macromolecular Interactions at High Resolution* 45–55 (Academic, New York, 1992).
21. Georgiadis, M. M. et al. *Structure* **3**, 879–892 (1995).
22. Jacobo-Molina, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6320–6324 (1993).
23. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. *Science* **256**, 1783–1790 (1992).
24. Brutlag, D. & Kornberg, A. *J. biol. Chem.* **247**, 241–248 (1972).
25. Joyce, C. M. *J. biol. Chem.* **264**, 10858–10866 (1989).
26. Otwinowski, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Burley, S.) 56–62 (SERC Daresbury Laboratory, UK, 1993).
27. Brünger, A. T. *X-PLOR version 3.1* (Yale Univ. Press, New Haven, CT, 1992).
28. Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
29. Lee, B. K. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
30. Tabor, S., Huber, H. E. & Richardson, C. C. *J. biol. Chem.* **262**, 16212–16223 (1987).

ACKNOWLEDGEMENTS. We thank Y. Kim (Yeungnam University, Korea) for assistance with the initiation of this project; D.-S. Lee (Korean Institute of Science and Technology) and S. W. Suh (Seoul National University) for the cDNA clone and partial supplement of the protein; R. M. Sweet for help with X12C beam line at the NSLS, Brookhaven National Laboratory; and J. Jäger, C. A. Brautigam and C. A. Joyce (Yale) for discussion. This work was supported by a grant from the American Cancer Society.

CORRESPONDENCE and requests for materials should be addressed to T.A.S. The Brookhaven Data Bank accession numbers for the *Taq*–DNA complex and revised apo structures and their structure factors are 1TAU and 1TAQ, and R1TAUSF and R1TAQSF, respectively.

KNOW YOUR COPY RIGHTS RESPECT OURS

The publication you are reading is protected by copyright law.
Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.
Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Guide to authors of contributions to *Nature*

Nature is a weekly international journal covering all the sciences. It is intended for an interdisciplinary readership, so all manuscripts should be written clearly and simply. Contributors should pay particular attention to the fact that English is not the first language of many readers. Space is limited and competition severe, so brevity is highly valued. The length limits below are specified in terms of the number of pages taken up in *Nature*. An uninterrupted page of text contains about 1,300 words, so a typical Letter to *Nature*, occupying 2½ pages, can contain about 1,500 words of text and four modest display items (figures and/or tables) with brief legends. There are no formal limits on the number of display items, but extra display material must be accompanied by commensurate reductions in the length of the text. Length limits can, however, be varied upwards or downwards at the editors' discretion.

Nature publishes the following types of contribution:

Review Articles survey recent developments in a topical area of scientific research. They should not generally occupy more than 6 pages of *Nature*.

Progress articles are relatively informal reviews of particularly topical and fast-moving fields. They do not exceed 4 pages of *Nature*.

Contributors wishing to submit a Review Article or Progress should first send a one-page synopsis to the Reviews Coordinator, from whom guidelines are available on request.

Articles are reports of original scientific research whose conclusions represent a substantial advance in understanding of an important problem and are of broad general interest. They do not generally exceed 5 pages of *Nature* and have no more than 50 references. They have a 'heading' of up to 80 words, often rewritten by *Nature's* editors, which advertises the content of the paper in general terms and should not contain numbers, abbreviations or measurements unless central to the message of the paper. The introduction and summary should be contained in the first three or four paragraphs of the main text. Layout of methods, figures and legends is summarized below.

Letters to *Nature* are short reports of original research focused on an outstanding finding whose implications are of interest to nonspecialists. They should not generally exceed 2½ pages of *Nature*. They have a bold first paragraph of not more than 180 words summarizing the background, rationale, conclusions and main implications of the work, aimed particularly at nonspecialist readers. Letters should have no more than 30 references. Layout of methods, figures and legends is summarized below.

Preparation of Articles and Letters

All manuscripts should be written in English, typed double-spaced on one side of the paper only. Manuscripts submitted in electronic form cannot be considered.

Manuscripts should be accompanied by a brief covering letter from the corresponding author, containing a full postal address, telephone and fax numbers, and e-mail address.

Five copies of manuscripts and original figures are needed, together with two copies of the covering letter. Five copies of manuscripts by any of the authors that have been submitted or are in press elsewhere should also be included, clearly marked as such; failure to do so may prejudice consideration of the manuscript.

Unless otherwise instructed by the editor handling a manuscript, when sending revised or resubmitted manuscripts five copies are required, each accompanied by a copy of the authors' response to referees' and editors' comments on the earlier version.

Titles should not exceed three lines of 31 characters each (including spaces), and should avoid numerical values, acronyms, abbreviations and punctuation wherever possible.

References should be numbered sequentially as they appear in the text, tables and figure legends. Only papers that have been published or submitted to a journal should be in the numbered list: manuscripts in preparation should be mentioned in the text with a list of authors. Reference lists should not contain textual material, grant details or acknowledgements.

Acknowledgements must be brief and appear after the reference list. *Nature* does not publish grant or contribution numbers.

Figures should not be larger than 22 by 28 cm unless unavoidable. They should not contain more than one panel unless the assembled parts are logically connected. Each copy should be marked with the corresponding author's name and, when known, the manuscript reference number. One photocopy of the original figures should be provided. All original figures are returned when *Nature* cannot offer to

publish a manuscript, but one copy of the manuscript and a photocopy of the figures are kept in confidential files for three months and then destroyed. Detailed instructions for figure preparation are available from the production department, including details of acceptable digital formats.

Colour figures are welcome but a contribution towards the cost of reproduction is requested. Inability to pay this charge will not prevent *Nature* publishing figures where colour is essential.

Tables. Symbols and abbreviations in a table should be defined immediately below the table, followed by essential descriptive material, all in double-spaced text. Tables should each be presented on a separate sheet of paper.

Methods. Brief details of materials and methods can be given in the text. Otherwise, they should be described at the end of the text in a section entitled "Methods", subdivided by bold headings referring to the methods used. This section should not exceed 800 words and ideally should be shorter. Additional material can be deposited on *Nature's* Web site. In contributions with methods sections, figure legends should not exceed 150 words each, and should be confined to a description of the figure and the symbols used.

Materials. As a condition of publication, authors are required to make materials and methods used freely available to academic researchers for their own use. Supporting datasets must be made available at the time of publication from the authors directly, by posting on *Nature's* Web site, by deposition in the appropriate public database or by distribution on the Internet. In the case of X-ray crystallographic coordinates only, public access may be delayed for up to one year after publication.

Articles and Letters should both contain a statement at the end of the text: "Correspondence and requests for materials to xxxx", with an e-mail address if desired, followed by the accession numbers or Web site address for supporting datasets. The designated author need not necessarily be the senior author, but *Nature* will expect the named author to handle matters arising from published contributions.

Manuscripts can be submitted to the Editor at *Nature*, Porter's South, Crinan St, London N1 9XW or at *Nature*, 968 National Press Building, Washington, DC 20045-1938. Proofs should be returned by express mail to London. Fax numbers and e-mail addresses for all editorial offices are on *Nature's* masthead.

Other contributions

Nature publishes informal contributed material as follows:

Correspondence contributions are short comments on topical issues, anecdotal material or reactions of readers to nontechnical material published in *Nature* (mainly News, Briefings, Opinion, Book Review or Commentary articles). All are unsolicited; submissions by e-mail to corres@nature.com are preferred (accents on names and places should be clearly indicated).

Commentary articles are informal, editorial comments on topical issues of public interest that have some bearing on scientific research. They are often controversial. Unsolicited manuscripts are considered, but enquiries can be made to the Commentary Editor by e-mail (nature@nature.com) or fax as a one-page synopsis or informal letter before formal submission.

News and Views articles inform nonspecialist readers about new scientific advances, either as reported in published papers or in the form of meeting reports. Most are commissioned, but proposals can be made to the News and Views Editor in advance of publication of the paper or of the meeting concerned. Authors are not allowed to discuss their own work or work from elsewhere in their institution. Detailed guidelines are available on request.

Scientific Correspondence is a predominantly peer-reviewed section of the journal specializing in very brief reports of general interest and discussion of topical issues. Priority is given to contributions of fewer than 500 words. Detailed guidelines are available on request. All contributions are unsolicited, and should be submitted by mail, not e-mail. Copies of contributions submitted by fax should be sent by mail with original figures and a cover letter noting that the manuscript has been faxed.

Book Reviews are all commissioned by the Book Review Editor. Unsolicited contributions are not considered.

Nature's editors will request final copy for Commentary, News and Views, Scientific Correspondence and Book Reviews on disk or via e-mail, as appropriate.

A more detailed Guide to Authors is available on *Nature's* WWW sites at <http://www.nature.com> or <http://www.america.nature.com>